

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 29, 2004, 23:24:18 ; Search time 46 Seconds
(without alignments)
1359,528 Million cell updates/sec

Title: US-10-057-832-2
Perfect score: 2072
Sequence: 1 MSREMDVDIAEVKPLVEKG.....TPNSGAGNSAGPKSEVSC 394

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq.19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2072	100.0	394	23	ABP52828
2	2072	100.0	394	24	ABU57628
3	2067	99.8	394	23	ABP65134
4	2067	99.8	394	24	ABR47543
5	2067	99.8	394	24	ABU07398
6	1250	60.3	375	23	ABU65166
7	1246	60.1	375	23	ABU65165
8	1231	59.4	363	22	ABG64392
9	1229	59.3	363	23	ABU65167

10	1224	59.1	363	22	AAW94019	Human stomach cano
11	1224	59.1	363	22	AAW94494	Human protein sequ
12	1158	55.9	359	23	ABU65168	Human NOV69d prote
13	1157	55.8	391	21	AAW32611	Human secreted pro
14	1137.5	54.9	339	22	AAW39870	Human polypeptide
15	1137.5	54.9	339	22	AAW47136	CDIF-17, Incyte I
16	1137.5	54.9	347	22	AAW41656	Human polypeptide
17	1102	53.2	251	21	AAW43199	Human cancer assoc
18	1049	50.6	371	23	AAU99505	Human Ndr2-related
19	1048	50.6	357	21	AAW41706	Human ORFX ORF1470
20	1048	50.6	357	23	ABW83808	Human P40. Homo s
21	1048	50.6	371	22	AAW79576	Human protein SEQ
22	1042	50.3	371	22	AAW78592	Human protein SEQ
23	1039	50.1	402	23	AAU99506	Human Ndr2-related
24	1039	50.1	575	23	AAU99507	Human Ndr2-related
25	936	45.2	337	23	ABW83809	Human P40-2. Homo
26	870	42.0	382	22	ABG10563	Novel human diagno
27	865	41.7	395	22	ABU31598	Novel human secret
28	859	41.5	299	23	ABU65169	Human NOV89e prote
29	855	41.3	325	23	ABW89674	Human polypeptide
30	850	41.0	247	22	ABG16537	Novel human diagno
31	812.5	39.2	286	22	ABW95462	Novel human diagno
32	731	35.3	251	22	ABG02995	Novel human diagno
33	601.5	29.0	343	22	ABW67375	Drosophila melanog
34	601.5	29.0	425	22	ABW70926	Drosophila melanog
35	601.5	29.0	429	22	ABW67374	Drosophila melanog
36	579.5	28.0	202	23	ABW89673	Novel human diagno
37	511	24.7	730	22	ABG19705	Novel human diagno
38	511	24.7	745	22	ABG02993	Novel human diagno
39	498.5	24.1	209	22	ABG10562	Novel human diagno
40	492.5	23.8	209	22	AAU31599	Novel human secret
41	464	22.4	87	21	AAU00716	Human secreted pro
42	402	19.4	106	23	ABG77040	Prostate specific
43	402	19.4	113	21	AAU00715	Human secreted pro
44	362	17.5	355	22	ABW88816	Drosophila melanog
45	356.5	17.2	347	21	AAW29136	Arabidopsis thalia

ALIGNMENTS

RESULT 1	ABP52828	ABP52828 standard; Protein; 394 AA.
XX	AC	ABP52828;
XX	DT	31-OCT-2002 (first entry)
XX	DE	Human CAP43 protein SEQ ID NO:2.
XX	KW	Human, CAP43; cytosolic; antiproliferative; cell targeted therapy; cancer; melanoma; lymphoma; malignant fibrous histiocytoma; tumour; metastasis; dysproliferative change; dysplasia; angiogenesis; psoriasis; ischaemia; atherosclerosis; stroke; inflammatory.
XX	OS	Homo sapiens.
XX	PN	W0200258719-A1.
XX	PD	01-AUG-2002.
XX	PF	25-JAN-2002; 2002MO-US01891.
XX	PR	25-JAN-2001; 2001US-264268P.
XX	PA	(UNY) UNIV NEW YORK STATE.
XX	PI	Costa M, Salnikow K, Yee H;
XX	DR	WPI; 2002-627386/67.
XX	DR	N-PSDB; AB074965.

PT Identifying a diseased cell or tissue, e.g. associated with abnormal
PT CAP43 expression, useful for diagnosing or detecting cancer, comprises
PT detecting an elevated level of CAP43 nucleic acid or gene product cell
PT or tissue

XX Claim 3, Fig 1B, 119pp; English.

XX The present sequence represents human CAP43. The present invention
CC describes a method (M1) for identifying a diseased cell or tissue, where
CC the disease is associated with abnormal CAP43 expression. M1 comprises
CC detecting an elevated level of CAP43 nucleic acid or gene product in a
CC sample of cells or tissues from the individual. CAP43 has cytoskeletal
CC and antiproliferative activities and can be used in cell targeted therapy.
CC The method is useful for identifying diseased cell or tissues,
CC specifically a disease associated with abnormal expression of CAP43,
CC such as cancer, e.g. lung, colon, kidney, breast, or prostate cancer,
CC melanoma, lymphoma, or malignant fibrous histiocytoma. Compounds, which
CC specifically bind to CAP43, are useful for treating cancer. Nucleic acids
CC comprising CAP43 fragments are useful as oligonucleotide probes and
CC primers to detect and amplify other nucleic acid molecules encoding
CC CAP43 analogues and homologues. CAP43 nucleic acids and proteins may
CC also be used in detection methods, and in treatment of tumours,
CC dysproliferative changes (e.g. metaplasias and dysplasias), disorders
CC involving inappropriate cell or tissue growth augmented by angiogenesis,
CC or psoriasis. These may further be used to identify atherosclerosis in
CC cells, tissues or organs, ischaemia, stroke, or inflammatory tissues.

XX Sequence 394 AA;

Query Match 100.0%; Score 2072; DB 23; Length 394;
Best Local Similarity 100.0%; Pred. No. 3.3e-194;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSREMODVLAELVKPLVEKGETITGLDQEPDVEODIETLHGSVHTLCGTPKGNRPVIL 60
DB 1 MSREMODVLAELVKPLVEKGETITGLDQEPDVEODIETLHGSVHTLCGTPKGNRPVIL 60
QY 61 TYHDIGNMHKTCTYNPLFNYEDMOEITQHFVAVCHVDAFGQDGAASFPAGYWPSPMDQLAE 120
DB 61 TYHDIGNMHKTCTYNPLFNYEDMOEITQHFVAVCHVDAFGQDGAASFPAGYWPSPMDQLAE 120
QY 121 MLPGVLAQDFGLKSIIGMGTAGAVILTRFALNPNPEWEGVILINVPNCAEGMDMAASKI 180
DB 121 MLPGVLAQDFGLKSIIGMGTAGAVILTRFALNPNPEWEGVILINVPNCAEGMDMAASKI 180
QY 121 MLPGVLAQDFGLKSIIGMGTAGAVILTRFALNPNPEWEGVILINVPNCAEGMDMAASKI 180
DB 121 MLPGVLAQDFGLKSIIGMGTAGAVILTRFALNPNPEWEGVILINVPNCAEGMDMAASKI 180
QY 181 SGMTOALPDWVSHLFGKEEMQSNVEVHTYRQHIYNDMNPGLHLFINAYNSRDLEIE 240
DB 181 SGMTOALPDWVSHLFGKEEMQSNVEVHTYRQHIYNDMNPGLHLFINAYNSRDLEIE 240
QY 181 SGMTOALPDWVSHLFGKEEMQSNVEVHTYRQHIYNDMNPGLHLFINAYNSRDLEIE 240
DB 181 SGMTOALPDWVSHLFGKEEMQSNVEVHTYRQHIYNDMNPGLHLFINAYNSRDLEIE 240
QY 241 RMPGTHHTVTLQCPALLVVDSSPAVDVVEGNSKLDPTKTTLLKMAACGGLPQISOPAK 300
DB 241 RMPGTHHTVTLQCPALLVVDSSPAVDVVEGNSKLDPTKTTLLKMAACGGLPQISOPAK 300
QY 301 LAEAFKYFVQGMGYMPSASMTRLMRSRTASGSSVTSLDGTRSRSHTSEGTSSHTSEGT 360
DB 301 LAEAFKYFVQGMGYMPSASMTRLMRSRTASGSSVTSLDGTRSRSHTSEGTSSHTSEGT 360
QY 361 RSRSHTSSEGAHLDTIPNSGAAGNSAGPKSMEVSC 394
DB 361 RSRSHTSSEGAHLDTIPNSGAAGNSAGPKSMEVSC 394

RESULT 2

ABUS7628
ID ABUS7628 standard; Protein; 394 AA.

XX AC ABUS7628;

DT 09-APR-2003 (first entry)

XX Differentially expressed breast cancer associated protein #15.

XX Breast cancer; differential gene expression; BC-cDNA;

KW breast cancer diagnosis; breast cancer monitoring;
KW breast cancer treatment; breast cancer staging.

OS Homo sapiens.

PN US2002156263-A1.

PD 24-OCT-2002.

PP 04-OCT-2001; 2001US-0974298.

PR 05-OCT-2000; 2000US-238331P.

PA (CHEN/) CHEN H.

PI Chen H;

DR WPI; 2003-182653/18.

XX New cDNAs, which are differentially expressed in (metastatic) breast
PT cancer useful for diagnosing or staging, breast cancer, or for
PT monitoring the treatment of breast cancer in an individual
XX Example; SEQ ID NO 81; 30pp; English.

CC The invention describes a combination of cDNAs (designated BC-cDNAs),
CC which are differentially expressed in breast cancer. The combination
CC includes 152 cDNA sequences, or their complements. The protein encoded
CC by any of these BC-cDNAs is useful for screening several molecules or
CC compounds to identify at least one ligand that specifically binds the
CC protein, producing or preparing polyclonal or monoclonal antibodies, or
CC purifying antibodies from a sample. The antibodies, which specifically
CC bind the protein differentially expressed in breast cancer is useful for
CC detecting the expression of a protein in a sample. The BC-cDNAs are
CC also useful for diagnosing, monitoring the treatment of, or staging,
CC breast cancer. This is the amino acid sequence of a differentially
CC expressed breast cancer associated protein.

CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from the US patent office at
CC seqdata.uspto.gov/sequence.html?DocID=20020156263.

XX Sequence 394 AA;

Query Match 100.0%; Score 2072; DB 24; Length 394;
Best Local Similarity 100.0%; Pred. No. 3.3e-194;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSREMODVLAELVKPLVEKGETITGLDQEPDVEODIETLHGSVHTLCGTPKGNRPVIL 60
DB 1 MSREMODVLAELVKPLVEKGETITGLDQEPDVEODIETLHGSVHTLCGTPKGNRPVIL 60
QY 61 TYHDIGNMHKTCTYNPLFNYEDMOEITQHFVAVCHVDAFGQDGAASFPAGYWPSPMDQLAE 120
DB 61 TYHDIGNMHKTCTYNPLFNYEDMOEITQHFVAVCHVDAFGQDGAASFPAGYWPSPMDQLAE 120
QY 121 MLPGVLAQDFGLKSIIGMGTAGAVILTRFALNPNPEWEGVILINVPNCAEGMDMAASKI 180
DB 121 MLPGVLAQDFGLKSIIGMGTAGAVILTRFALNPNPEWEGVILINVPNCAEGMDMAASKI 180
QY 121 MLPGVLAQDFGLKSIIGMGTAGAVILTRFALNPNPEWEGVILINVPNCAEGMDMAASKI 180
DB 121 MLPGVLAQDFGLKSIIGMGTAGAVILTRFALNPNPEWEGVILINVPNCAEGMDMAASKI 180
QY 181 SGMTOALPDWVSHLFGKEEMQSNVEVHTYRQHIYNDMNPGLHLFINAYNSRDLEIE 240
DB 181 SGMTOALPDWVSHLFGKEEMQSNVEVHTYRQHIYNDMNPGLHLFINAYNSRDLEIE 240
QY 181 SGMTOALPDWVSHLFGKEEMQSNVEVHTYRQHIYNDMNPGLHLFINAYNSRDLEIE 240
DB 181 SGMTOALPDWVSHLFGKEEMQSNVEVHTYRQHIYNDMNPGLHLFINAYNSRDLEIE 240
QY 241 RMPGTHHTVTLQCPALLVVDSSPAVDVVEGNSKLDPTKTTLLKMAACGGLPQISOPAK 300
DB 241 RMPGTHHTVTLQCPALLVVDSSPAVDVVEGNSKLDPTKTTLLKMAACGGLPQISOPAK 300
QY 301 LAEAFKYFVQGMGYMPSASMTRLMRSRTASGSSVTSLDGTRSRSHTSEGTSSHTSEGT 360
DB 301 LAEAFKYFVQGMGYMPSASMTRLMRSRTASGSSVTSLDGTRSRSHTSEGTSSHTSEGT 360
QY 361 RSRSHTSSEGAHLDTIPNSGAAGNSAGPKSMEVSC 394

Db 361 RSRSHTEGAHLDTTPNSGAAGNSAGPKSMSEVSC 394

RESULT 3

ABP65134 ID ABP65134 standard; Protein; 394 AA.

AC ABP65134;

DT 12-NOV-2002 (first entry)

DE Hypoxia-regulated protein #8.

XX Cyclostatic; vasoregulator; antiatherosclerotic; gene therapy;
 XX antiinflammatory; vulnary; gynecological; ophthalmological; vaccine;
 XX hypoxia; tumorigenesis; angiogenesis; apoptosis; cancer;
 XX ischaemic condition; reperfusion injury; retinopathy; neonatal stress;
 XX preclapemia; atherosclerosis; inflammatory condition; wound healing;
 XX inflammation; erythropoiesis; hair loss; human.

OS Homo sapiens.

PN WO200246465-A2.

PD 13-JUN-2002.

PE 10-DEC-2001; 2001WO-GB05458.

PR 08-DEC-2000; 2000GB-0030076.

PR 08-FEB-2001; 2001GB-0003156.

PR 25-OCT-2001; 2001GB-0025666.

PA (OXFORD) OXFORD BIOMEDICA UK LTD.

PI White J, Mundy CR, Ward NR, Krige D, Kingsman SM, Harris RA;
 PI Rayner WN;

DR WPI; 2002-627238/67.

XX Identifying a gene involved in disease for treating hypoxia-regulated
 PT conditions, comprises comparing the transcriptome/proteome of two cell
 PT types under different conditions and identifying a differentially
 PT regulated gene -

PS Claim 35; Page 335; 538pp; English.

XX The present invention relates to methods for identifying genes and
 CC proteins that are implicated in a specific disease or physiological
 CC condition. The method comprises comparing the transcriptome/proteome of a
 CC specialised cell type implicated in a disease or condition with that of a
 CC second specialised cell type, under two experimental conditions, and
 CC identifying a gene that is differentially regulated in the two
 CC specialised cell types under experimental conditions. ABV7873-ABV78116
 CC and ABP65061-ABP65257 were identified using the methods of the invention.
 CC The coding sequences and proteins are useful for treating a disease in a
 CC patient, for manufacture of a medicament for treating hypoxia-regulated
 CC conditions, and for regulating tumorigenesis, angiogenesis, apoptosis,
 CC biological response to hypoxia conditions, or hypoxic-associated
 CC pathology in a patient. The coding sequences and proteins are also useful
 CC for monitoring the therapeutic treatment of a disease or physiological
 CC condition, such as cancer, ischaemic conditions, reperfusion injury,
 CC retinopathy, neonatal stress, preclapemia, atherosclerosis, inflammatory
 CC conditions, wound healing, inflammation, erythropoiesis or hair loss.

XX Sequence 394 AA;

Query Match 99.8%; Score 2067; DB 23; Length 394;

Best Local Similarity 99.7%; Pred No. 1e-193;

Matches 393; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSREMODVLAELVKELVEKGETITGLQEPDVQEDITLHGSVHTLCGTPKGNRPVIL 60

Db 1 MSREMODVLAELVKELVEKGETITGLQEPDVQEDITLHGSVHTLCGTPKGNRPVIL 60
 QY TYHDIGNHKTQVPLNFVEDMOBITQHPAVCHVDAPGQDGAASPPAGVYVPSMDQAE 120
 Db 61 TYHDIGNHKTQVPLNFVEDMOBITQHPAVCHVDAPGQDGAASPPAGVYVPSMDQAE 120
 QY 121 MLPGLVQFGKLSIIGGTGAGAYILTRFALNDEMEVGLIVNPFCAEGMDMAASKI 180
 Db 121 MLPGLVQFGKLSIIGGTGAGAYILTRFALNDEMEVGLIVNPFCAEGMDMAASKI 180
 QY 181 SGWTQALPDVWVSHLFKEKEQSNVEVHTYRQHTVNDNPNGLHLFINAYSRRDLEIE 240
 Db 181 SGWTQALPDVWVSHLFKEKEQSNVEVHTYRQHTVNDNPNGLHLFINAYSRRDLEIE 240
 QY 241 RPPMGHTVTLTQCPALIVGDSPPAVDAVVECNKLPPTKTTLLKMDCCGLPQISOPAK 300
 Db 241 RPPMGHTVTLTQCPALIVGDSPPAVDAVVECNKLPPTKTTLLKMDCCGLPQISOPAK 300
 QY 301 LAEAFKTFVQGMGYPASMTRLMRSRTASGSSVTSLDGTRSRSHTEGTRSRSHTEGT 360
 Db 301 LAEAFKTFVQGMGYPASMTRLMRSRTASGSSVTSLDGTRSRSHTEGTRSRSHTEGT 360
 QY 361 RSRSHTEGAHLDTTPNSGAAGNSAGPKSMSEVSC 394
 Db 361 RSRSHTEGAHLDTTPNSGAAGNSAGPKSMSEVSC 394

RESULT 4

ABR47543 ID ABR47543 standard; Protein; 394 AA.

AC ABR47543;

DT 12-JUN-2003 (first entry)

DE Breast cancer associated protein sequence SEQ ID NO:324.

XX Human; breast cancer; cytostatic; gene therapy.

OS Homo sapiens.

PN WO2003004989-A2.

PD 16-JAN-2003.

PE 21-JUN-2002; 2002WO-US19669.

PR 21-JUN-2001; 2001US-299887P.

PR 27-JUN-2001; 2001US-301572P.

PR 18-JUL-2001; 2001US-306501P.

PR 25-SEP-2001; 2001US-325002P.

PR 05-MAR-2002; 2002US-362585P.

PR 14-MAY-2002; 2002US-380391P.

PA (MILL-) MILLENITUM PHARM INC.

PI Lillie J, Gannavarapu M, Glat K, Hoerh S, Kamatkar S, Mertens M;
 PI Monahan JB, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE, Bast RC;
 PI Hototobagyi GN, Puzestai L, Meric F, Sahin A, Mills GB;

DR WPI; 2003-210381/20.

DR N-PSDB; ACC50241.

PT Breast cancer diagnosis or treatment by comparing the level of
 PT expression of a marker in a patient sample with that in the control
 PT non-breast cancer sample -

PS Claim 1; SEQ ID 324; 128pp; English.

XX The present invention describes a method for assessing whether a patient
 CC is afflicted with breast cancer. The method comprises comparing the level
 CC of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
 CC ABR47386 to ABR47632) in a patient sample and the normal level of

expression of the marker in a control non-breast cancer sample, where a significant increase in the level of expression of the marker in the patient sample and the normal level is an indication that the patient is afflicted with breast cancer. The breast cancer associated sequences from the present invention have cytosolic activities and can be used in gene therapy. The method is useful for diagnosing and treating breast cancer.

N.B. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPRO at ftp.wipro.int/pub/published_pct_sequences.

Sequence 394 AA:

Query Match 99.8%; Score 2067; DB 24; Length 394;
Best Local Similarity 99.7%; Pred. No. 1e-193;
Matches 393; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 MSREMDDVDLAEVKPLVEKGETITGLQEPDVOEDIEIHLGSHVHTLCGTPKGNRPVIL 60
1 MSREMDDVDLAEVKPLVEKGETITGLQEPDVOEDIEIHLGSHVHTLCGTPKGNRPVIL 60
61 TYHDIGMNHKTCTNPLFNEDMOEITQHPAVCHVDAPGQDGAASFPAGMYPSMDQLAE 120
61 TYHDIGMNHKTCTNPLFNEDMOEITQHPAVCHVDAPGQDGAASFPAGMYPSMDQLAE 120
121 MLPGLVLOQFGLKSIIGMGTGAGAYILTRPALNPNPEWEGVLINVPCEAGMDMAASKI 180
121 MLPGLVLOQFGLKSIIGMGTGAGAYILTRPALNPNPEWEGVLINVPCEAGMDMAASKI 180
181 SGMTOALPDMVVSHLFGKEKMOQSNVEVHTYRQHIIVNDMPGNLHFINAYNSRRDLEIE 240
181 SGMTOALPDMVVSHLFGKEKMOQSNVEVHTYRQHIIVNDMPGNLHFINAYNSRRDLEIE 240
241 RMPGTHHTVTLQCPALLVVDSSPAVDVAVECNSKLDPTKTTLLKMAOCGGLPOISQPAK 300
241 RMPGTHHTVTLQCPALLVVDSSPAVDVAVECNSKLDPTKTTLLKMAOCGGLPOISQPAK 300
301 LAEAKRYFVQGMGYMPSASMTRLMRSRTASGSVTSLDGTRSRSHTSEGTSSRSHTSEGT 360
301 LAEAKRYFVQGMGYMPSASMTRLMRSRTASGSVTSLDGTRSRSHTSEGTSSRSHTSEGT 360
361 RSRSHTSBGALHDITPNSGAAGNSAGPXSMEVSC 394
361 RSRSHTSBGALHDITPNSGAAGNSAGPXSMEVSC 394

RESULT 5
ABU07398 standard; Protein; 394 AA.

ABU07398
AC ABU07398;

28-JAN-2003 (first entry)

Protein differentially regulated in prostate cancer #1.

Prostate cancer; gene expression; differential regulation;
molecular marker; drug target; cancer detection; cancer diagnosis;
cancer staging; cancer grading; cancer assessment; cancer monitoring.

Homo sapiens.

WO200281638-A2.

17-OCT-2002.

08-APR-2002; 2002WO-US10824.

06-APR-2001; 2001US-281731P.
06-APR-2001; 2001US-281732P.

(ORIG-) ORIGENE TECHNOLOGIES INC.

PI Sun Z, Jay G;
XX WPI; 2003-058520/05.

Novel genes which are differentially regulated in prostate cancer,
useful for diagnosing prostate cancer in prostate tissue sample and
assessing therapeutic or preventive intervention in prostate cancer
patients -

Claim 1; Page 198-199; 41sep; English.

The invention describes genes (I) which are differentially regulated in prostate cancer. (I) is useful for diagnosing a prostate cancer in a sample comprising prostate tissue, which involves determining the number of target genes which are differentially-regulated in the sample, where the number is indicative of the probability that the sample comprises prostate cancer. (I) is useful for assessing a therapeutic or preventive intervention in a subject having a prostate cancer, which involves determining the expression levels in a sample comprising prostate tissue of target genes which are differentially-regulated in prostate cancer. Preferably, the expression levels of at least 10 genes are determined. (I) is also useful for identifying agents that modulate a biological activity of a polypeptide differentially-regulated in prostate cancer cells, which involves contacting a polypeptide differentially-regulated in prostate cancer cells with a test agent under conditions effective for determining whether the test agent modulates the biological activity. (I) is useful as molecular markers, as drug targets, and for detecting, diagnosing, staging, grading, assessing, monitoring, prognosticating, preventing or treating, determining predisposition to diseases and conditions especially relating to prostate cancer. (I) and its expression products are used in the diagnostic test to assay for presence of cancer e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in blood etc. (I) is useful for assessing cancer e.g., to determine the type of cancer, its stage of development, the nature of genetic defect, etc. The polypeptide encoded by (I) can be used as target for therapy or drug discovery. (I) can also be used for expressing the polypeptide and thus for searching specific binding partners of the polypeptide. (I) is useful in therapeutic applications to treat prostate cancer. The identification of specific genes, and groups of genes, expressed in pathways physiologically relevant to prostate cancer delineates the definition of functional and disease pathways and the delineation of targets in these pathways which are useful in diagnostic, therapeutic, and clinical applications. This is the amino acid sequence of a protein differentially regulated in prostate cancer.

Sequence 394 AA:

Query Match 99.8%; Score 2067; DB 24; Length 394;
Best Local Similarity 99.7%; Pred. No. 1e-193;
Matches 393; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 MSREMDDVDLAEVKPLVEKGETITGLQEPDVOEDIEIHLGSHVHTLCGTPKGNRPVIL 60
1 MSREMDDVDLAEVKPLVEKGETITGLQEPDVOEDIEIHLGSHVHTLCGTPKGNRPVIL 60
61 TYHDIGMNHKTCTNPLFNEDMOEITQHPAVCHVDAPGQDGAASFPAGMYPSMDQLAE 120
61 TYHDIGMNHKTCTNPLFNEDMOEITQHPAVCHVDAPGQDGAASFPAGMYPSMDQLAE 120
121 MLPGLVLOQFGLKSIIGMGTGAGAYILTRPALNPNPEWEGVLINVPCEAGMDMAASKI 180
121 MLPGLVLOQFGLKSIIGMGTGAGAYILTRPALNPNPEWEGVLINVPCEAGMDMAASKI 180
181 SGMTOALPDMVVSHLFGKEKMOQSNVEVHTYRQHIIVNDMPGNLHFINAYNSRRDLEIE 240
181 SGMTOALPDMVVSHLFGKEKMOQSNVEVHTYRQHIIVNDMPGNLHFINAYNSRRDLEIE 240
241 RMPGTHHTVTLQCPALLVVDSSPAVDVAVECNSKLDPTKTTLLKMAOCGGLPOISQPAK 300
241 RMPGTHHTVTLQCPALLVVDSSPAVDVAVECNSKLDPTKTTLLKMAOCGGLPOISQPAK 300
301 LAEAKRYFVQGMGYMPSASMTRLMRSRTASGSVTSLDGTRSRSHTSEGTSSRSHTSEGT 360

```

Db      301 LAEAFKYVQCGWGPMSASMTRLMRSRTVSGSSTSLDGTSSRSHTSEGTSSRSHTS 360
QY      361 RSRHTSEGAHLDITPNNGAAGNSAGPKSMVEVSC 394
Db      361 RSRHTSEGAHLDITPNNGAAGNSAGPKSMVEVSC 394

RESULT 6
ABU65166
ID      ABU65166 standard; Protein; 375 AA.
XX
AC      ABU65166;
XX
DT      20-MAY-2003 (first entry)
XX
DE      Human NOV98b protein.
XX
KW      NOXV; cytosstatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
KW      hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
KW      human.
XX
OS      Homo sapiens.
XX
PN      WO200272757-A2.
XX
PD      19-SEP-2002.
XX
PE      08-MAR-2002; 2002WO-US06908.
XX
PR      08-MAR-2001; 2001US-274101P.
PR      08-MAR-2001; 2001US-274194P.
PR      08-MAR-2001; 2001US-274281P.
PR      08-MAR-2001; 2001US-274322P.
PR      09-MAR-2001; 2001US-274849P.
PR      12-MAR-2001; 2001US-275335P.
PR      13-MAR-2001; 2001US-275578P.
PR      13-MAR-2001; 2001US-275579P.
PR      13-MAR-2001; 2001US-275601P.
PR      14-MAR-2001; 2001US-276000P.
PR      16-MAR-2001; 2001US-276776P.
PR      19-MAR-2001; 2001US-276994P.
PR      20-MAR-2001; 2001US-277239P.
PR      20-MAR-2001; 2001US-277321P.
PR      20-MAR-2001; 2001US-277327P.
PR      21-MAR-2001; 2001US-277791P.
PR      22-MAR-2001; 2001US-277833P.
PR      23-MAR-2001; 2001US-278152P.
PR      26-MAR-2001; 2001US-278894P.
PR      27-MAR-2001; 2001US-278999P.
PR      27-MAR-2001; 2001US-279036P.
PR      28-MAR-2001; 2001US-279344P.
PR      30-MAR-2001; 2001US-277388P.
PR      30-MAR-2001; 2001US-279995P.
PR      30-MAR-2001; 2001US-280233P.
PR      02-APR-2001; 2001US-280802P.
PR      02-APR-2001; 2001US-280822P.
PR      02-APR-2001; 2001US-280900P.
PR      04-APR-2001; 2001US-281194P.
PR      13-APR-2001; 2001US-283675P.
PR      30-APR-2001; 2001US-287424P.
PR      02-MAY-2001; 2001US-288066P.
PR      03-MAY-2001; 2001US-288342P.
PR      03-MAY-2001; 2001US-288528P.
PR      15-MAY-2001; 2001US-291190P.
PR      16-MAY-2001; 2001US-291099P.
PR      16-MAY-2001; 2001US-291240P.
PR      30-MAY-2001; 2001US-294485P.
PR      31-MAY-2001; 2001US-294899P.
PR      18-JUN-2001; 2001US-299027P.
PR      19-JUN-2001; 2001US-299303P.
PR      19-JUN-2001; 2001US-299310P.

```

```

PR      10-JUL-2001; 2001US-304354P.
PR      31-JUL-2001; 2001US-309198P.
PR      16-AUG-2001; 2001US-312903P.
PR      10-SEP-2001; 2001US-318463P.
PR      12-SEP-2001; 2001US-318770P.
PR      27-SEP-2001; 2001US-325430P.
PR      27-SEP-2001; 2001US-325681P.
PR      18-OCT-2001; 2001US-330380P.
PR      31-OCT-2001; 2001US-335301P.
PR      14-NOV-2001; 2001US-333172P.
PR      14-NOV-2001; 2001US-333271P.
PR      14-NOV-2001; 2001US-333272P.
PR      14-NOV-2001; 2001US-333184P.
PR      14-NOV-2001; 2001US-333272P.
PR      21-NOV-2001; 2001US-333094P.
PR      21-NOV-2001; 2001US-337426P.
PR      03-DEC-2001; 2001US-338092P.
PR      04-DEC-2001; 2001US-337185P.
PR      03-JAN-2002; 2002US-345705P.
PR      07-MAR-2002; 2002US-0092900.
XX
XX
PA      (CURA-) CURAGEN CORP.
XX
PI      Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CBA, Li L,
PI      Zethusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
PI      Paturajan M, Gangoli E, Vernet CAM, Guo X, Tchernev V;
PI      Fernandes ER, Casman SJ, Malvankar UM, Gerlach V, Liu Y;
PI      Anderson D, Spaderna SK, Carterton E, Burgess C, Lette M, Zhong H;
PI      Allobrook JP, Lepley DM, Rieger DK;
XX
XX      WPI; 2002-723332/78.
DR      N-PSDB; ABX97133.
XX
PT      NOXV polypeptides and polynucleotides, useful for preventing or
PT      treating a disorder associated with aberrant NOXV expression or
PT      activity e.g., cancer, hypertension, atherosclerosis, cardiomyopathy or
PT      bronchial asthma.
XX
XX      Claim 1; Page 353; 1103jp; English.
XX
XX      This invention describes novel human NOXV polypeptides which have
XX      cytosstatic, cardiant, antiarteriosclerotic, antiasthmatic and
XX      hypotensive activity. Pharmaceutical compositions comprising the NOXV
XX      proteins or nucleic acid molecules or NOXV antibodies are useful for
XX      preventing or treating a disorder associated with aberrant NOXV
XX      expression or activity e.g. cancer, hypertension, atherosclerosis,
XX      cardiomyopathy or bronchial asthma. The products of the invention can
XX      be used for gene therapy or in a vaccine. ABU65041-ABU65218 represent
XX      the NOXV polypeptides encoded by ABX97008-ABX97185.
XX
XX      Sequence 375 AA:
XX
XX      Query Match 60.3%; Score 1250; DB 23; Length 375;
XX      Best Local Similarity 61.5%; Pred. No. 1,26-113;
XX      Matches 243; Conservative 59; Mismatches 67; Indels 26; Gaps 6;
QY      4 EMOVDLAENVKEKKEITGLLOEPDVQODIETLHGSVHTLCGPKGNRPVILTYH 63
Db      3 ELQDVQLTEIKPL-LNDKNGTRNFQDPCCQDHEITLHGVHTVIRGLPKGNRPVILTYH 61
QY      64 DIGMNHKTQVNPFLNVEDMOEITQHPAVCHVADAGOODGASFPAGVYPPMDLAEMLP 123
Db      62 DIGMNHKSCFPAFNFEMOEBITQHPAVCHVADAGOODGASFPPTGYOYPMDELAETLP 121
QY      124 GVLQDFGLKSTIGKGTGAGAYILTRPALNDEWVEGLVINVPQABGMDMAASKISGW 183
Db      122 PVLTLTLSTKSTIGGVGAGAYILSRFALNHRQLVEGLVILNVDCAKMDMAASKISGL 181
QY      184 TQALPDWVSHLPGKEKESQVVEVHTYROIIVNDMPGNLHLPLNANSRDIEIRPM 243
Db      182 TTNVVDIILAHNFGQEBLQALDLIOTYRMTIADINDNDQLPLNSNGRRDIEIRPT 241
QY      244 PGTH--TVTLQCPALLVVDSSPAVDAVVECNKSLDPTKTLTKMADCGGLPOISOPAK 300

```

```

Db      242 LQONNKSKTLKCSSTLLVVDNSPAVEAVEGNSRLNPITLLKWAQGGILPQVVOPEK 301
Oy      301 LAEAFKYPQGMWPSASMTLRMSRT-AGSSVTSJUDGTSRSHTEGTSRSHTEG 359
Db      302 LLEAFKYPQGMWPSASMTLRMSRTSHSTSSJGSGSPPSRSVT-----SNQSDG 354
Oy      360 TRSRSHTEGAHLDTTPNSGAAGNSAGPKSMVEVC 394
Db      355 TQESCESP-----DVLDRH-----QTMVEVC 375

RESULT 7
ABU65165
ID      ABU65165 standard; Protein; 375 AA.
XX
AC      ABU65165;
XX
DT      20-MAY-2003 (first entry)
XX
DE      Human NOV89a protein.
XX
KM      NOX; cytosolic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
KW      hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
XX      human.
XX
OS      Homo sapiens.
XX
PN      WO200272757-A2.
XX
PD      19-SEP-2002.
XX
PF      08-MAR-2002; 2002WO-US06908.
XX
PR      08-MAR-2001; 2001US-274101P.
PR      08-MAR-2001; 2001US-274194P.
PR      08-MAR-2001; 2001US-274281P.
PR      08-MAR-2001; 2001US-274322P.
PR      09-MAR-2001; 2001US-274849P.
PR      12-MAR-2001; 2001US-275235P.
PR      13-MAR-2001; 2001US-275578P.
PR      13-MAR-2001; 2001US-275579P.
PR      13-MAR-2001; 2001US-275601P.
PR      14-MAR-2001; 2001US-276000P.
PR      16-MAR-2001; 2001US-276776P.
PR      19-MAR-2001; 2001US-276994P.
PR      20-MAR-2001; 2001US-277239P.
PR      20-MAR-2001; 2001US-277321P.
PR      21-MAR-2001; 2001US-277791P.
PR      22-MAR-2001; 2001US-277833P.
PR      23-MAR-2001; 2001US-278152P.
PR      26-MAR-2001; 2001US-278894P.
PR      27-MAR-2001; 2001US-278999P.
PR      27-MAR-2001; 2001US-279034P.
PR      28-MAR-2001; 2001US-279346P.
PR      30-MAR-2001; 2001US-277338P.
PR      30-MAR-2001; 2001US-279995P.
PR      30-MAR-2001; 2001US-280233P.
PR      02-APR-2001; 2001US-280802P.
PR      02-APR-2001; 2001US-280822P.
PR      04-APR-2001; 2001US-280900P.
PR      13-APR-2001; 2001US-281194P.
PR      13-APR-2001; 2001US-283675P.
PR      30-APR-2001; 2001US-287424P.
PR      02-MAY-2001; 2001US-288066P.
PR      03-MAY-2001; 2001US-288342P.
PR      15-MAY-2001; 2001US-291190P.
PR      16-MAY-2001; 2001US-291099P.
PR      16-MAY-2001; 2001US-291240P.
PR      30-MAY-2001; 2001US-294485P.
PR      31-MAY-2001; 2001US-294489P.

```

```

PR      31-MAY-2001; 2001US-294899P.
PR      18-JUN-2001; 2001US-299027P.
PR      19-JUN-2001; 2001US-299303P.
PR      19-JUN-2001; 2001US-299310P.
PR      10-JUL-2001; 2001US-304354P.
PR      31-JUL-2001; 2001US-309198P.
PR      16-AUG-2001; 2001US-312903P.
PR      10-SEP-2001; 2001US-318462P.
PR      12-SEP-2001; 2001US-318770P.
PR      27-SEP-2001; 2001US-325430P.
PR      27-SEP-2001; 2001US-325681P.
PR      18-OCT-2001; 2001US-330380P.
PR      31-OCT-2001; 2001US-335301P.
PR      14-NOV-2001; 2001US-332172P.
PR      14-NOV-2001; 2001US-332271P.
PR      14-NOV-2001; 2001US-332272P.
PR      14-NOV-2001; 2001US-333184P.
PR      14-NOV-2001; 2001US-333272P.
PR      21-NOV-2001; 2001US-332094P.
PR      03-DEC-2001; 2001US-337426P.
PR      03-DEC-2001; 2001US-338092P.
PR      04-DEC-2001; 2001US-337185P.
PR      03-JAN-2002; 2002US-345705P.
PR      07-MAR-2002; 2002US-0092900.
XX
PA      (CURA-) CURAGEN CORP.
XX
PI      Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L,
PI      Zernusen BD, Gusev V, Ji W, Gorman L, Miller CB, Kekuda R,
PI      Paturajan W, Gangoli E, Vernet CM, Guo X, Tchernev V,
PI      Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y,
PI      Anderson D, Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H,
PI      Alabrook JP, Lepley DM, Rieger DK,
XX
XX      WPI, 2002-723332/78.
DR      N-PSDB; ABX97132.
XX
XX      NOX polypeptides and polynucleotides, useful for preventing or
PT      treating a disorder associated with aberrant NOX expression or
PT      activity e.g., cancer, hypertension, atherosclerosis, cardiomyopathy or
PT      bronchial asthma -
XX
XX      Claim 1; Page 352; 1103pp; English.
XX
XX      This invention describes novel human NOX polypeptides which have
CC      cytosolic, cardiant, antiarteriosclerotic, antiasthmatic and
CC      hypotensive activity. Pharmaceutical compositions comprising the NOX
CC      proteins or nucleic acid molecules or NOX antibodies are useful for
CC      preventing or treating a disorder associated with aberrant NOX
CC      expression or activity e.g. cancer, hypertension, atherosclerosis,
CC      cardiomyopathy or bronchial asthma. The products of the invention can
CC      be used for gene therapy or in a vaccine. ABU65041-ABU65218 represent
CC      the NOX polypeptides encoded by ABX97008-ABX97185.
XX
SQ      Sequence 375 AA;
Oy      4 EMQDVDAEVRKPLVERGETITGLQEFVQEDIEITLHGSVHTLCTGPKGNRPVILTYH 63
Db      3 ELQDVQLTEIKPLL-NDKNGTRNFQDFDCQEHDIETTHGVVHTTIGLPRGNRPVILTYH 61
Oy      64 DIGMNHKTCYNPLFNTEDNDEITQHPAVCHVDAPGQDGAASPAGYMPSPMDLAEMLP 123
Db      62 DIGLNKSCFNAPFNEDMDEITQHPAVCHVDAPGQDGAASPPTQYPTMDLEMLP 121
Oy      124 GVLQFGKLSIIGMGAGAGYILTRPALNPEWVEGLVILNVPACAGMWDMAASKISG 183
Db      122 PVLTHLSLSIIGIVGAGAYILSRPALNPELVESGLVILNVPCKAGMWDMAASKISGL 181
Oy      184 TQALPDMVVSHLFGKEEMQSNVEVHTYRQHIIVDNKPGNHLIFINAYNSRDLIERPM 243

```

```

Db      182 TTNVVDIIIAHFGQELQANLDIQYRMHIAODINODNLQFLNNGRDLIERPI 241
Qy      244 PGTH---TVTLQCPALIVGDSSPAVDVAVECNSKLDPTKTLTKMADCCGLPOISOPAK 300
Db      242 LGQDNKSKTKLKSTLLVVGDNSPAWEAVVECNRLNPINTLLKMDCCGLPVVOPGK 301
Qy      301 LAEAFKYFVQGMGYMPASMTRLMRSRT-ASGSSVTSLDGTRSRSHTSSEGTGRSRSHTSSEG 359
Db      302 LTFAPKYFLQGMGYVPSASMTRLMRSRTHSTSSLSGSGESPFRSVT-----SNQSDG 354
Qy      360 TRSRSHTSSEGAHLDTTPNSGAAGSAGPKSMVEVSC 394
Db      355 TQESCESP-----DVLDRH-----QTMVEVSC 375

RESULT 8
AAG64392
ID      AAG64392 standard; Protein; 363 AA.
XX
AC      AAG64392;
XX
DT      04-OCT-2001 (first entry)
XX
DE      Human reducing agent and tunicamycin-responsive protein 40.
XX
KW      Human; reducing agent; tunicamycin-responsive protein 40; cytostatic;
KW      virucidal; immunomodulatory; antiinflammatory; haemostatic;
KW      malignant tumour; haemopathy; HIV infection; immunological disease;
KW      inflammation.
XX
OS      Homo sapiens.
XX
PN      WO200155375-A1.
XX
PD      02-AUG-2001.
XX
PF      15-JAN-2001; 2001WO-CN00050.
XX
PR      26-JAN-2000; 2000CN-011517.
XX
PA      (BIOD-) BIODOR GENE TECHNOLOGY LTD SHANGHAI.
PI      Mao Y, Xie Y.
XX
DR      WPI; 2001-488795/53.
XX
DR      N-PSDB; AAH73967.
XX
PT      Human reducing agent and tunicamycin-responsive protein 40 and encoded
PT      polynucleotide, applicable in diagnosis and treatment of malignant
PT      tumor, hemopathy, immunological diseases and various inflammations -
XX
PS      Claim 1; Page 36; 40pp; Chinese.
XX
CC      The present sequence is the protein sequence for human reducing agent and
CC      tunicamycin-responsive protein 40. The protein and its coding sequence
CC      are useful in the diagnosis and treatment of malignant tumour,
CC      haemopathy, HIV infection, immunological diseases and various
CC      inflammations.
XX
SQ      Sequence 363 AA;

Query Match      59.4%; Score 1231; DB 22; Length 363;
Best Local Similarity 60.5%; Pred. No. 8,6e-112;
Matches 239; Conservative 57; Mismatches 61; Indels 38; Gaps 6;

Qy      4 EMQDVDAEVKPLVKEGETITGLQEPVQODITFLGVSIVTLCGTPKGNRPVILTYH 63
Db      3 ELQDVQLIRIKLND-----KEHIDITTHGVAVHTIRGLPKGNRPVILTYH 49

Qy      64 DIGMNHKTCTYNPLFNYEDMQEITQHPAVCHVDAPGQDGAASFPAGYMPSPMDLAEMLP 123
Db      50 DIGLNHKSCTFNAPFNFEDMQEITQHPAVCHVDAPGQDGAASFPAGYMPSPMDLAEMLP 109

```

```

Qy      124 GVLQGFLEKSIIGWGTGAGAYILTRFALNPNPEWEGVLINVPDCAEGMDMAASKISGW 183
Db      110 SVLTHLSKSIIGVGAGAYILSRFALNHPELVEGLVINVPDCAGMDMAASKISGL 169
Qy      184 TQALPDVAVSHLFGKEEMQSNVEVHTYRQHIYNDMPGUILFYNASRRDLIERPM 243
Db      170 TTNVVDIIIAHFGQELQANLDIQYRMHIAODINODNLQFLNNGRDLIERPI 229
Qy      244 PGTH---TVTLQCPALIVGDSSPAVDVAVECNSKLDPTKTLTKMADCCGLPOISOPAK 300
Db      230 LGQDNKSKTKLKSTLLVVGDNSPAWEAVVECNRLNPINTLLKMDCCGLPVVOPGK 289
Qy      301 LAEAFKYFVQGMGYMPASMTRLMRSRT-ASGSSVTSLDGTRSRSHTSSEGTGRSRSHTSSEG 359
Db      290 LTFAPKYFLQGMGYVPSASMTRLMRSRTHSTSSLSGSGESPFRSVT-----SNQSDG 342
Qy      360 TRSRSHTSSEGAHLDTTPNSGAAGSAGPKSMVEVSC 394
Db      343 TQESCESP-----DVLDRH-----QTMVEVSC 363

RESULT 9
ABU65167
ID      ABU65167 standard; Protein; 363 AA.
XX
AC      ABU65167;
XX
DT      20-MAY-2003 (first entry)
XX
DE      Human NOV89c protein.
XX
KW      NOVX; cytostatic; cardiac; antiarteriosclerotic; antiasthmatic; cancer;
KW      hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
KW      human.
XX
OS      Homo sapiens.
XX
PN      WO200272757-A2.
XX
PD      19-SEP-2002.
XX
PF      08-MAR-2002; 2002WO-US06908.
XX
PR      08-MAR-2001; 2001US-274101P.
PR      08-MAR-2001; 2001US-274194P.
PR      08-MAR-2001; 2001US-274281P.
PR      08-MAR-2001; 2001US-274322P.
PR      09-MAR-2001; 2001US-274849P.
PR      12-MAR-2001; 2001US-275235P.
PR      13-MAR-2001; 2001US-275578P.
PR      13-MAR-2001; 2001US-275579P.
PR      13-MAR-2001; 2001US-275799P.
PR      13-MAR-2001; 2001US-276001P.
PR      14-MAR-2001; 2001US-276000P.
PR      16-MAR-2001; 2001US-27676P.
PR      19-MAR-2001; 2001US-276994P.
PR      20-MAR-2001; 2001US-277239P.
PR      20-MAR-2001; 2001US-277321P.
PR      20-MAR-2001; 2001US-277327P.
PR      20-MAR-2001; 2001US-277791P.
PR      21-MAR-2001; 2001US-277833P.
PR      22-MAR-2001; 2001US-277833P.
PR      23-MAR-2001; 2001US-278152P.
PR      26-MAR-2001; 2001US-278894P.
PR      27-MAR-2001; 2001US-278999P.
PR      27-MAR-2001; 2001US-279036P.
PR      28-MAR-2001; 2001US-279344P.
PR      30-MAR-2001; 2001US-277338P.
PR      30-MAR-2001; 2001US-279995P.
PR      30-MAR-2001; 2001US-280233P.
PR      02-APR-2001; 2001US-280802P.
PR      02-APR-2001; 2001US-280822P.
PR      02-APR-2001; 2001US-280900P.
PR      04-APR-2001; 2001US-281194P.

```

13-APR-2001; 2001US-283675P.
 PR 30-APR-2001; 2001US-287424P.
 PR 02-MAY-2001; 2001US-288066P.
 PR 03-MAY-2001; 2001US-288342P.
 PR 03-MAY-2001; 2001US-288528P.
 PR 15-MAY-2001; 2001US-291190P.
 PR 16-MAY-2001; 2001US-291099P.
 PR 16-MAY-2001; 2001US-291240P.
 PR 30-MAY-2001; 2001US-294485P.
 PR 31-MAY-2001; 2001US-294899P.
 PR 31-MAY-2001; 2001US-294899P.
 PR 18-JUN-2001; 2001US-299027P.
 PR 19-JUN-2001; 2001US-299303P.
 PR 19-JUN-2001; 2001US-299310P.
 PR 10-JUL-2001; 2001US-304354P.
 PR 31-JUL-2001; 2001US-309198P.
 PR 16-AUG-2001; 2001US-312903P.
 PR 10-SEP-2001; 2001US-318462P.
 PR 12-SEP-2001; 2001US-318770P.
 PR 27-SEP-2001; 2001US-325430P.
 PR 27-SEP-2001; 2001US-325681P.
 PR 18-OCT-2001; 2001US-330380P.
 PR 31-OCT-2001; 2001US-335301P.
 PR 14-NOV-2001; 2001US-332172P.
 PR 14-NOV-2001; 2001US-332271P.
 PR 14-NOV-2001; 2001US-332372P.
 PR 14-NOV-2001; 2001US-333184P.
 PR 14-NOV-2001; 2001US-333272P.
 PR 21-NOV-2001; 2001US-332094P.
 PR 03-DEC-2001; 2001US-337426P.
 PR 03-DEC-2001; 2001US-338092P.
 PR 04-DEC-2001; 2001US-337185P.
 PR 03-JAN-2002; 2002US-345705P.
 PR 07-MAR-2002; 2002US-0092900.
 (CURA-) CURAGEN CORP.

PA Padigar M, Spytek KA, Shenoy SG, Taupier RJ, Pena CRA, Li L;
 PI Zernusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
 PI Paturajan M, Gangoli E, Vermet CAM, Guo X, Tchenev V;
 PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y;
 PI Anderson D, Spaderma SK, Carteron E, Burgess C, Lette M, Zhong H;
 PI Alsobrook JP, Lepley DM, Rieger DK;
 DR MPI: 2002-723332/78.
 DR N-PSDB; ABX97134.

XX NOVA polypeptides and polynucleotides, useful for preventing or
 PT treating a disorder associated with aberrant NOVA expression or
 PT activity e.g., cancer, hypertension, atherosclerosis, cardiomyopathy or
 PT bronchial asthma -
 XX Claim 1; Page 353; 1103pp; English.

XX This invention describes novel human NOVA polypeptides which have
 CC cytosolic, cardiant, antiarteriosclerotic, antiasthmatic and
 CC hypotensive activity. Pharmaceutical compositions comprising the NOVA
 CC proteins or nucleic acid molecules or NOVA antibodies are useful for
 CC preventing or treating a disorder associated with aberrant NOVA
 CC expression or activity e.g. cancer, hypertension, atherosclerosis,
 CC cardiomyopathy or bronchial asthma. The products of the invention can
 CC be used for gene therapy or in a vaccine. ABU5041-ABU6518 represent
 CC the NOVA polypeptides encoded by ABX97008-ABX97185.

XX Sequence 363 AA;

Query Match 59.3%; Score 1229; DB 23; Length 363;
 Best Local Similarity 60.5%; Pred. No. 1.3e-111; Indels 38; Gaps 6;
 Matches 239; Conservative 57; Mismatches 61;

QY 4 EMQVDLAELVRLVKEGKITGLQEPDVGQDITLTLGSHVHTLCKGPKGNRPVILTYH 63
 DB 3 ELQDVQLREIKRLAND-----KEHDIETHTGVVHTIRGLPKGNRPVILTYH 49

QY 64 DIGMHHKTCVNPLEPNVEDMOEITQHPAVCHVDAPGQODGASAPAGYMPMDQLAEMLP 123
 DB 50 DIGMHHKTCVNPLEPNVEDMOEITQHPAVCHVDAPGQODGASAPAGYMPMDQLAEMLP 109
 QY 124 GVLQOFGIKSIITGNGAGAYITLRLPANNPEWEGVLVINVPCEAGMDMAASKISGW 183
 DB 110 PVLTHLSKSIIGIGVAGAYILSRFALNHPVEGLVINVPCEAGMDMAASKISGL 169
 QY 184 TQALPDVAVSHLREKEMQSNVEVHTYRQIYVNDMPGULHPIFANYNRRDLEIFRM 243
 DB 170 TTVNVVDIIIAHFFQOEELQANLDIQTFRMHQIDQDMLOFLNSYNRRDLEIFRPI 229
 QY 244 PGTH---TVTLQCPALIVGDSSPAVDPAVCEKSLDPTTTLTKMADCGGLPOLISQPAK 300
 DB 230 LGQNDKSKTKLCKSTLVLVGDSPAVAVCEKSLRNPITTLTKMADCGGLPOLVQDPK 289
 QY 301 LAEAFKYFVGGMGMPASMTRLRSRT-ASGSSVTLSDGTRSRSHTSSEGTSRSHTSSEG 359
 DB 290 LTFEAFKYFVGGMGMPASMTRLRSRTHTSSLSGSESPFERSVT-----SNQSDG 342
 QY 360 TRSRSHTSSEGAHLDITNCSAGNSAGPKSMVSC 394
 DB 343 TQESCESP-----DVLDRH-----QTMVEVSC 363

RESULT 10

AA094019 standard; Protein; 363 AA.

AA094019;

13-NOV-2001 (first entry)

Human stomach cancer expressed polypeptide SEQ ID NO 108.

Human; stomach cancer; marker; screening; micro-metastasis;

peritoneal dissemination.

Homo sapiens.

WC020109317-A1.

08-FEB-2001.

28-JUL-2000; 2000MO-JP05063.

29-JUL-1999; 99JP-0248036.

27-AUG-1999; 99JP-0300253.

18-OCT-1999; 99US-0159590.

11-JAN-2000; 2000JP-0118776.

17-FEB-2000; 2000US-0183322.

02-MAY-2000; 2000JP-0183767.

09-JUN-2000; 2000JP-0241899.

WPI: 2001-570287/64.

N-PSDB; AA193896.

(HELI-) HELIX RES INST.

Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T, Aburatani H;
 PI Kodama T, Midorikawa Y;

XX Claim 1; Page 189-190; 242pp; Japanese.

XX The invention relates to stomach cancer-expressed genes
 CC (AA193842-AA193917) and the encoded proteins (AA093967-AA094039). The
 CC genes can be used as markers in blood tests for screening for the early
 CC stages of the disease. The proteins and peptides can be used as targets

CC for screening for compounds to treat the disease. They can also be used
CC for predicting micro-metastases. The gene can predict peritoneal
CC dissemination.

XX Sequence 363 AA:

Query Match 59.1%; Score 1224; DB 22; Length 363;
Best Local Similarity 60.5%; Pred. No. 4.2e-111;
Matches 239; Conservative 56; Mismatches 62; Indels 38; Gaps 6;

QY 4 EMQVDVLAEVKPLVEKGETITGLQEPDVOQEDIEITLHGSVHTLCGTPKGNRPVILTYH 63
DB 3 ELQDVQLTEIKPLND-----KEDIETTHGVVHTIRGLPKGNRPVILTYH 49
QY 64 DIGMNHKTCYNPLFNYEDMOEITOHFAVCHVDAPQODGASFPAGVYPSMDOLAEMLP 123
DB 50 DIGLNHKSNSAFNPFEDMOEITOHFAVCHVDAPQODGASFPAGVYPTMDELAEMLP 109
QY 124 GYLQDFGLKSIIGMGTGAGAYILTRFALNPEMVEGLVLINVPACAEQMDMAASKISGM 183
DB 110 PVLTHLSLKSITIGIGVAGAYILSRFALNHPVLVEGLVLINVDPCAKGMDMAASKISGL 169
QY 184 TQALPDVMSHLFGKEBMSQNVVHTTRQHIYVNDMPGNLHLPFINANSRRDLEIERPM 243
DB 170 TTNVVDITIAHFGQEBELQANLDLQTYRMHIAODINQDNLQFLINSYNGRRDLEIERPI 229
QY 244 PGTH---TVTLQCPALLVGDSSPAVDVAVECNKSLDPTKTLTKMADCGGLPQVVOGPK 300
DB 230 LGQNDNKSCTLKCSLTLLVGDNSPAVEAVNCSRLNPINTLLKMDCCGLPQVVOGPK 289
QY 301 LAEAFKYFVQMGVMPASMTLRMSRT-ASGSSVTSIDGTRSRSHTEGTRSRSHTEG 359
DB 290 LTEAFKFLQMGVYIPASMTLRASRTHSTSSLSGSESPFSRSVT-----SNQSDG 342
QY 360 TRSRSHTEGAHLDITPNSGAAGSAGPKSMVEVC 394
DB 343 TOESCESP-----DVLDRH-----QTMVEVC 363

RESULT 11

AA894494
ID AAB94494 standard; Protein; 363 AA.

XX AAB94494;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:15186.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

PN EPI074617-A2.

PD 07-FEB-2001.

PE 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Iehi S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI, 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602

PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -

XX Claim 8, SEQ ID 15186; 2537bp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification, where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the 5' end of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX Sequence 363 AA:

Query Match 59.1%; Score 1224; DB 22; Length 363;
Best Local Similarity 60.5%; Pred. No. 4.2e-111;
Matches 239; Conservative 56; Mismatches 62; Indels 38; Gaps 6;

QY 4 EMQVDVLAEVKPLVEKGETITGLQEPDVOQEDIEITLHGSVHTLCGTPKGNRPVILTYH 63
DB 3 ELQDVQLTEIKPLND-----KEDIETTHGVVHTIRGLPKGNRPVILTYH 49
QY 64 DIGMNHKTCYNPLFNYEDMOEITOHFAVCHVDAPQODGASFPAGVYPSMDOLAEMLP 123
DB 50 DIGLNHKSNSAFNPFEDMOEITOHFAVCHVDAPQODGASFPAGVYPTMDELAEMLP 109
QY 124 GYLQDFGLKSIIGMGTGAGAYILTRFALNPEMVEGLVLINVPACAEQMDMAASKISGM 183
DB 110 PVLTHLSLKSITIGIGVAGAYILSRFALNHPVLVEGLVLINVDPCAKGMDMAASKISGL 169
QY 184 TQALPDVMSHLFGKEBMSQNVVHTTRQHIYVNDMPGNLHLPFINANSRRDLEIERPM 243
DB 170 TTNVVDITIAHFGQEBELQANLDLQTYRMHIAODINQDNLQFLINSYNGRRDLEIERPI 229
QY 244 PGTH---TVTLQCPALLVGDSSPAVDVAVECNKSLDPTKTLTKMADCGGLPQVVOGPK 300
DB 230 LGQNDNKSCTLKCSLTLLVGDNSPAVEAVNCSRLNPINTLLKMDCCGLPQVVOGPK 289
QY 301 LAEAFKYFVQMGVMPASMTLRMSRT-ASGSSVTSIDGTRSRSHTEGTRSRSHTEG 359
DB 290 LTEAFKFLQMGVYIPASMTLRASRTHSTSSLSGSESPFSRSVT-----SNQSDG 342
QY 360 TRSRSHTEGAHLDITPNSGAAGSAGPKSMVEVC 394
DB 343 TOESCESP-----DVLDRH-----QTMVEVC 363

RESULT 12

ABU65168
ID ABU65168 standard; Protein; 359 AA.

XX ABU65168;

DT 20-MAY-2003 (first entry)

DE Human NOV99d protein.

XX NOX; cytosolic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
KW hypertensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
KW human.
XX Homo sapiens.
XX MO200272757-A2.
XX 19-SEP-2002.
XX 08-MAR-2002; 2002WO-US06908.
XX 08-MAR-2001; 2001US-274101P.
XX 08-MAR-2001; 2001US-274194P.
XX 08-MAR-2001; 2001US-274281P.
XX 08-MAR-2001; 2001US-274322P.
XX 09-MAR-2001; 2001US-274849P.
XX 12-MAR-2001; 2001US-275235P.
XX 13-MAR-2001; 2001US-275578P.
XX 13-MAR-2001; 2001US-275579P.
XX 13-MAR-2001; 2001US-275601P.
XX 14-MAR-2001; 2001US-276000P.
XX 16-MAR-2001; 2001US-276766P.
XX 19-MAR-2001; 2001US-276994P.
XX 20-MAR-2001; 2001US-277239P.
XX 20-MAR-2001; 2001US-277321P.
XX 20-MAR-2001; 2001US-277327P.
XX 21-MAR-2001; 2001US-277791P.
XX 22-MAR-2001; 2001US-277833P.
XX 23-MAR-2001; 2001US-278152P.
XX 26-MAR-2001; 2001US-278894P.
XX 27-MAR-2001; 2001US-278999P.
XX 27-MAR-2001; 2001US-279036P.
XX 28-MAR-2001; 2001US-279344P.
XX 30-MAR-2001; 2001US-279388P.
XX 30-MAR-2001; 2001US-279958P.
XX 30-MAR-2001; 2001US-280233P.
XX 02-APR-2001; 2001US-280802P.
XX 02-APR-2001; 2001US-280822P.
XX 02-APR-2001; 2001US-280900P.
XX 04-APR-2001; 2001US-281194P.
XX 13-APR-2001; 2001US-283675P.
XX 30-APR-2001; 2001US-287424P.
XX 02-MAY-2001; 2001US-288066P.
XX 03-MAY-2001; 2001US-288343P.
XX 03-MAY-2001; 2001US-288528P.
XX 15-MAY-2001; 2001US-291190P.
XX 16-MAY-2001; 2001US-291099P.
XX 16-MAY-2001; 2001US-291240P.
XX 30-MAY-2001; 2001US-294485P.
XX 31-MAY-2001; 2001US-294899P.
XX 31-MAY-2001; 2001US-294899P.
XX 18-JUN-2001; 2001US-299027P.
XX 19-JUN-2001; 2001US-299303P.
XX 19-JUN-2001; 2001US-299310P.
XX 10-JUL-2001; 2001US-304354P.
XX 31-JUL-2001; 2001US-309198P.
XX 16-AUG-2001; 2001US-312903P.
XX 10-SEP-2001; 2001US-318462P.
XX 12-SEP-2001; 2001US-318770P.
XX 27-SEP-2001; 2001US-325430P.
XX 27-SEP-2001; 2001US-325681P.
XX 18-OCT-2001; 2001US-330380P.
XX 31-OCT-2001; 2001US-333011P.
XX 14-NOV-2001; 2001US-332172P.
XX 14-NOV-2001; 2001US-332717P.
XX 14-NOV-2001; 2001US-332722P.
XX 14-NOV-2001; 2001US-333184P.
XX 14-NOV-2001; 2001US-333272P.
XX 21-NOV-2001; 2001US-333094P.
XX 03-DEC-2001; 2001US-337426P.
XX 03-DEC-2001; 2001US-338092P.

PR 04-DEC-2001; 2001US-337185P.
PR 03-JAN-2002; 2002US-345705P.
PR 07-MAR-2002; 2002US-0092900.
XX
XX
XX
XX
XX (CURA-) CURAGEN CORP.
XX Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L,
PI Zethusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R,
PI Patturajan M, Gangoli E, Vernet CAM, Guo X, Tchernev V,
PI Fernandes ER, Caeman SJ, Malyankar UM, Gerlach V, Liu Y,
PI Anderson D, Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H,
PI Alsobrook JP, Lepley DM, Rieger DK,
XX
XX WPI; 2002-723332/78.
XX N-PSDB; ABX97135.
XX
XX NOX polypeptides and polynucleotides, useful for preventing or
PT treating a disorder associated with aberrant NOX expression or
PT activity e.g., cancer, hypertension, atherosclerosis, cardiomyopathy or
PT bronchial asthma
XX
XX Claim 1, Page 354; 1103pp; English.
XX
XX This invention describes novel human NOX polypeptides which have
CC cytosolic, cardiant, antiarteriosclerotic, antiasthmatic and
CC hypotensive activity. Pharmaceutical compositions comprising the NOX
CC proteins or nucleic acid molecules or NOX antibodies are useful for
CC preventing or treating a disorder associated with aberrant NOX
CC expression or activity e.g. cancer, hypertension, atherosclerosis,
CC cardiomyopathy or bronchial asthma. The products of the invention can
CC be used for gene therapy or in a vaccine. AB065041-AB065218 represent
CC the NOX polypeptides encoded by ABX97008-ABX97185.
XX
XX Sequence 359 AA;
SQ
Query Match 55.9%; Score 1158; DB 23; Length 359;
Best Local Similarity 58.2%; Pred. No. 1,2e-104;
Matches 230; Conservative 55; Mismatches 68; Indels 42; Gaps 7;
QY 4 EMOVDVLAIEYKPLVEKGETITGLQDFDVOEDIEFLHGSVHTLCTPRKGRPVILTYH 63
DB 3 ELQDVQLTEIKPL-NDKNGTRNFQDFDCEHDIETTHGVVHTINGLPKGRPVILTYH 61
QY 64 DIAMNKTCTNPLFNFEDMOEITQHFVAVGVDAFGQDGAAPFAGWMPSPMDQLAMLP 123
DB 62 DIGANKSCFNAPFNFEDMOEITQHFVAVGVDAFGQDGAAPFAGWMPSPMDQLAMLP 121
QY 124 GVLQDFGLKSIIGMGTGAGAYITLTPALNNPEWEGVLINVPCEAGMDWAASKISGM 183
DB 122 PVLTHLSLKSIIGIGVAGAYILSRFALNHPLEVEGLVINVPCEAGMDWAASKISGL 181
QY 184 TQALPDMVYSHLFGKEMQSNVEVHTYQHYNDMPNGNHLFIYAANSRRDLETERPM 243
DB 182 TTVVVDIILAHFGQBELQANLDLQTYRMHIAQDINDONLQFLNYSYNGRRDLETERPI 241
QY 244 PGTH--TWTLQCPALVYVGDSSPAVDVVEGSKLDPKRTTLKADCGGLPQISQPAK 300
DB 242 LGQNDKSKTKLKSTLLVGDNSPAVEAV-----MADCGGLPQVQPAK 285
QY 301 LAEAFKYPVQMGVMPASNTLRMSRT-ASGSSVYSLDGTSSRSHTSRSHTSBG 359
DB 286 FTAEAFKYPVQMGVMPASNTLRMSRTHSTSSLSGSSPFSRSVY-----SNQSDG 338
QY 360 TTSRSHTSBGALDITPNSGAAGNSAGPSPSMVSC 394
DB 339 TQSCESP-----DVLDRH-----QTMVESC 359
RESULT 13
AAB23611
ID AAB23611 standard; Protein; 391 AA.
XX
AC AAB23611;

```

XX 12-JAN-2001 (first entry)
XX
XX Human secreted protein SEQ ID NO: 22.
XX
XX Human; secreted protein; cytokine; cell proliferation;
XX nutritional supplement; immune modulation; autoimmune disorder;
XX haematopoiesis regulation; tissue growth; haemostasis; inflammation.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 10..22
XX Protein /label= signal_peptide
XX /label= mature_protein
XX
XX WO200049134-A1.
XX
XX 24-AUG-2000.
XX
XX 18-FEB-2000; 2000WO-US04340.
XX
XX 19-FEB-1999; 99US-0120680.
XX 23-APR-1999; 99US-0298733.
XX 17-AUG-1999; 99US-0149639.
XX 23-SEP-1999; 99US-0155686.
XX 01-OCT-1999; 99US-0157247.
XX 29-NOV-1999; 99US-0167822.
XX 29-NOV-1999; 99US-0167823.
XX 15-FEB-2000; 2000US-0298733.
XX
XX (ALPH-) ALPHAGENE INC.
XX
XX Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;
XX
XX MPI; 2000-549267/50.
XX N-PSDB; AAA93111.
XX
XX New secreted proteins and polynucleotides encoding them, which are
XX derived from Homo sapiens, useful for therapy, diagnosis, and research,
XX as well as nutritional sources or supplements -
XX
XX Claim 31; Page 255-257; 309pp; English.
XX
XX The present sequence is the sequence of a human secreted protein. Its
XX cDNA was isolated from an adult pancreas cDNA library. The proteins
XX and coding sequences of the invention can be used in the isolation of
XX similar genes and proteins, in the elucidation of their function in vivo,
XX and to treat a number of conditions. It is possible that they may have
XX uses as nutritional supplements, as cytokine or cell proliferation
XX factors, in immune modulation, where they may be used to treat immune and
XX autoimmune diseases, as haematopoiesis regulators (treating myeloid or
XX lymphoid cell deficiencies), in the promotion of tissue growth, they may
XX have chemokinetic or chemotactic activity, haemostatic or thrombolytic
XX activity, or anti-inflammatory activity.
XX
XX Sequence 391 AA:
XX
XX Query Match 55.8%; Score 1157; DB 21; Length 391;
XX Best Local Similarity 56.3%; Pred. No. 1.7e-104;
XX Matches 224; Conservative 58; Mismatches 78; Indels 38; Gaps 4;
XX
XX 5 MODVLAAYKPLVEGETTG-----LLOEPVQVQODITTLGASVHTLCGPKGNR 56
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 24 LQELRPFEPKPL-RGDQATBLESSDAFLADTWMKRDIDTPYGLHVVIRGSPKGNR 82
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 57 PILTYHDIQGMNHCYVNPENYEDMOETTHFAVCHVDAPGQDGAASFPAGWYPSMD 116
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 83 PAIILYHDVGLNHCYNTFFNFDQETTKAFVCHVDAPQDQASQFPGSYGPPSWH 142
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 117 QLAEMLPGLVQOFGKSLIGMGTGAGAVLTTRPALNPEMVEGLVLINVPACAEQMMDWA 176
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

DB 143 QLAAMLPVVOHFGFKVYIGVGAGAYVLAKFALLIPDLVEGLVLVINIDNGKMWIDWA 202
QY ASKISGWTQALPDMVSHLPGKEKMOSSVEVHTYRQIYVDNKPNGNHLPTINAYNSRD 236
DB 203 ATKLSGLSTLPDLVHLFSQSELVNNTLEVOYRQIGVNVQANLQLEWNNNSRD 262
QY 237 LEIERPMPGTHVTTLQCPALLVGDSPPADAVVECNKSLDPTKTYLLKMACDGLPQIS 296
DB 263 LDINRPQVPAKTLRCPVNLVVGDNAPAEQGVVECNKSLDPTTTLKMACDGLPQVT 322
QY 297 QPAKLARAFKTFVQGMPSASMTRLMRSTASGSVTSIDGTRNSHTSEGTRNSHT 356
DB 323 QPGKLTAFKTFVQGMPSASMTRLMRSTASGSVTSIDGTRNSHTSEGTRNSHT 376
QY 357 SEGTRNSHTSEGAHLDITPNSGAAGSAGPKSKMEVSC 394
DB 377 SEGIGQVNHNT-----NEVSC 391

RESULT 14
AAM39870
ID AAM39870 standard; Protein; 339 AA.
XX
XX AAM39870;
XX
XX 22-OCT-2001 (first entry)
XX
XX Human polypeptide SEQ ID NO 3015.
XX
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX leukaemia.
XX
XX Homo sapiens.
XX
XX WO200153312-A1.
XX
XX 26-JUL-2001.
XX
XX 26-DEC-2000; 2000WO-US34263.
XX
XX 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0598042.
XX 19-JUL-2000; 2000US-0620312.
XX 03-AUG-2000; 2000US-0653450.
XX 14-SEP-2000; 2000US-0662191.
XX 19-OCT-2000; 2000US-0693036.
XX 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YF, Liu C, Aundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX Zhao QH, Zhou P, Goodrich R, Drmanac RT;
XX
XX MPI; 2001-442253/47.
XX N-PSDB; AAI59026.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
XX
XX Example 4; SEQ ID NO 3015; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
XX the encoded polypeptides (AAM38642-AA44213) with nootropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and

```



```

Db 128 PDLYEGLVLVNIIDPNGKGMIDMAATKLSGLTSTLPDVTLSHLFSQEBELVNNTELVSQYRQ 187
Qy 214 HIYNDMPGNLHLFINAYNSRRDLIERPMPGTHVTLOCPALVVGDSPPAVDAVEGN 273
Db 188 QIGNVVNOANLQLFMMMTNSRDLIDINRPGTVFPAKTLRCPVMLVVGDNAPADGVVEGN 247
Qy 274 SKLDPTKTTLLKMACCGGLPQISQPAKLAFAKYFVQGWGYPSPASMTRLMRSRTASGSS 333
Db 248 SKLDPTTTTFLKMACDSGGLPVOTPGKLTFAKCYFLOGWGYPSPASMTRLASRTASLTS 307
Qy 334 VTSLDCTSRSHSTSECTSRSHSTSECTSRSHSTSEGAHLDTIPNSGAAGNSAGPKSMEVS 393
Db 308 ASSVDGSRPQACT-----HSESSEGLAQVNH-----MEVS 338
Qy 394 C 394
Db 339 C 339

```

Search completed: January 29, 2004, 23:27:54
 Job time : 48 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 29, 2004, 23:29:24 ; Search time 40 Seconds
(without alignments)
2047.067 Million cell updates/sec

Title: US-10-057-832-2

Perfect score: 2072
Sequence: 1 MSRMQVDLAEVKPLVEKKG.....TPNSGAAGNSAGPSKMEVSC 394

Scoring table:

BLASTN62
Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

Published Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pdp:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pdp:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pdp:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pdp:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pdp:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pdp:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pdp:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pdp:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pdp:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pdp:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pdp:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pdp:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pdp:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pdp:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pdp:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pdp:*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pdp:*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pdp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2072	100.0	394	US-09-974-298-81	Sequence 81, Appl
2	2072	100.0	394	US-09-919-039-197	Sequence 197, Appl
3	2072	100.0	394	US-10-172-069-33	Sequence 33, Appl
4	2072	100.0	394	US-10-057-832-2	Sequence 2, Appl
5	2067	99.8	394	US-10-177-283-324	Sequence 324, Appl
6	1960	94.6	394	US-10-205-194-156	Sequence 156, Appl
7	1817	87.7	356	US-10-108-260A-3803	Sequence 3803, Appl
8	1646.5	79.5	323	US-10-108-260A-3274	Sequence 3274, Appl
9	1137.5	54.9	339	US-10-094-749-1842	Sequence 1842, Appl
10	1102	53.2	251	US-09-925-301-1404	Sequence 1404, Appl
11	1049	50.6	371	US-10-172-069-1	Sequence 1, Appl
12	1039	50.1	402	US-10-172-069-2	Sequence 2, Appl
13	1038	50.1	371	US-10-172-069-32	Sequence 32, Appl
14	892.5	43.1	311	US-10-104-047-2056	Sequence 2056, Appl
15	855	41.3	325	US-10-264-237-2050	Sequence 2050, Appl

16	579.5	28.0	202	US-10-264-237-2049	Sequence 2049, Appl
17	402	19.4	106	US-10-000-256A-146	Sequence 146, Appl
18	313	15.1	112	US-09-867-550-1070	Sequence 1070, Appl
19	219	10.6	339	US-09-864-761-35439	Sequence 35439, Appl
20	169	8.2	36	US-09-864-761-36332	Sequence 36332, Appl
21	113.5	5.5	314	US-10-214-473-74	Sequence 74, Appl
22	113.5	5.5	314	US-10-272-490-74	Sequence 74, Appl
23	113	5.5	34	US-09-864-761-34925	Sequence 34925, Appl
24	107.5	5.2	303	US-10-156-761-8055	Sequence 8055, Appl
25	106	5.1	19723	US-10-084-846A-5	Sequence 5, Appl
26	104	5.0	266	US-10-156-761-9087	Sequence 9087, Appl
27	104	5.0	500	US-10-369-493-20424	Sequence 20424, Appl
28	104	5.0	537	US-10-321-802-2	Sequence 2, Appl
29	101.5	4.9	252	US-10-369-493-424	Sequence 424, Appl
30	100	4.8	300	US-10-156-761-8083	Sequence 8083, Appl
31	99	4.8	305	US-10-080-644-11	Sequence 11, Appl
32	98.5	4.8	262	US-09-903-410-38	Sequence 38, Appl
33	98.5	4.8	262	US-10-027-805-38	Sequence 38, Appl
34	98.5	4.8	262	US-10-027-804-38	Sequence 38, Appl
35	98.5	4.8	270	US-10-080-644-9	Sequence 9, Appl
36	98	4.7	374	US-09-738-626-3961	Sequence 3961, Appl
37	98	4.7	476	US-09-070-844-26	Sequence 26, Appl
38	98	4.7	487	US-09-070-844-24	Sequence 24, Appl
39	98	4.7	512	US-09-070-844-4	Sequence 4, Appl
40	98	4.7	526	US-09-070-844-2	Sequence 2, Appl
41	98	4.7	966	US-09-978-698-2	Sequence 2, Appl
42	97.5	4.7	224	US-10-099-322-303	Sequence 303, Appl
43	97	4.7	307	US-10-156-761-11388	Sequence 11388, Appl
44	97	4.7	536	US-09-930-218-2	Sequence 2, Appl
45	97	4.7	536	US-10-431-438-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1	US-09-974-298-81
Sequence 81, Application US/09974298	
Patent No. US20020156263A1	
GENERAL INFORMATION:	
APPLICANT: Chen, Huei-Mei	
FILE REFERENCE: PA-0037 P	
CURRENT FILING DATE: 2001-10-04	
CURRENT FILING DATE: 2001-10-04	
PRIOR APPLICATION NUMBER: 60/238,331	
PRIOR FILING DATE: 2000-05-10	
NUMBER OF SEQ ID NOS: 194	
SOFTWARE: PERL Program	
SEQ ID NO 81	
LENGTH: 394	
TYPE: PRT	
ORGANISM: Homo sapiens	
FEATURE:	
NAME/KEY: misc feature	
OTHER INFORMATION: Incyte ID No. US20020156263A1 1969731CD1	
US-09-974-298-81	
Query Match	100.0%; Score 2072; DB 10; Length 394;
Best Local Similarity	100.0%; Pred. No. 5.2e-199;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MSRMQVDLAEVKPLVEKETTIGLQERDVQODLETILHGSVHTLCGPKKNRPVIL 60
DB	1 MSRMQVDLAEVKPLVEKETTIGLQERDVQODLETILHGSVHTLCGPKKNRPVIL 60
QY	TYHDIIGNHHTCTVPLFNVEDMOEITQHFVAVCHVDADGQODGASPPAGVYVPSMDOLAE 120
DB	TYHDIIGNHHTCTVPLFNVEDMOEITQHFVAVCHVDADGQODGASPPAGVYVPSMDOLAE 120
QY	MLPGLVLOQFGSLIIGGTGAGAYILTRFALNBNBEGVTLINVNCAEGMDMAASKI 180
DB	MLPGLVLOQFGSLIIGGTGAGAYILTRFALNBNBEGVTLINVNCAEGMDMAASKI 180

; CURRENT FILING DATE: 2002-01-25
 ; PRIOR APPLICATION NUMBER: US 10/057,832
 ; PRIOR FILING DATE: 2002-01-25
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 2
 ; LENGTH: 394
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-057-832-2

Query Match 100.0%; Score 2072; DB 14; Length 394;
 Best Local Similarity 100.0%; Pred. No. 5,26-199;
 Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSREMODVLAELVKRLVEKGETITGLLOEPDVEQODIETLHGSVHTLCTGPKGNRPVIL 60
 DB 1 MSREMODVLAELVKRLVEKGETITGLLOEPDVEQODIETLHGSVHTLCTGPKGNRPVIL 60
 QY 61 TYHDIGMNHKTCYNPLRYEEDMOEITQHFAVCHVDAPGQDGAASFPAGYMPSPMDOLAE 120
 DB 61 TYHDIGMNHKTCYNPLRYEEDMOEITQHFAVCHVDAPGQDGAASFPAGYMPSPMDOLAE 120
 QY 121 MLPGLVLOQFGLKSIIGMGTGAGAYILTRFALNPEMEVGLVLIIVNPCAEGMMDWAASKI 180
 DB 121 MLPGLVLOQFGLKSIIGMGTGAGAYILTRFALNPEMEVGLVLIIVNPCAEGMMDWAASKI 180
 QY 181 SGWTOALPDVNVSHLFGKEEMQSNVEVHTYRQHIIVDMNPGNLHLFINAVNSRRDLEIE 240
 DB 181 SGWTOALPDVNVSHLFGKEEMQSNVEVHTYRQHIIVDMNPGNLHLFINAVNSRRDLEIE 240
 QY 241 RPEMGTHTVTLQCPALLVVGDSPPAVDAVVECNKSLDPTKTTLLKMDCCGLPOLISQPAK 300
 DB 241 RPEMGTHTVTLQCPALLVVGDSPPAVDAVVECNKSLDPTKTTLLKMDCCGLPOLISQPAK 300
 QY 301 LAEAKFYVQGMGMPASMTRLMRSRTASGSSVTSLDGTRSRSHTSGTRSRSHTSGT 360
 DB 301 LAEAKFYVQGMGMPASMTRLMRSRTASGSSVTSLDGTRSRSHTSGTRSRSHTSGT 360
 QY 361 RSRHTSEGAHLDTTPNSGAAGNSAGPKSMEVSC 394
 DB 361 RSRHTSEGAHLDTTPNSGAAGNSAGPKSMEVSC 394

RESULT 5
 ; Sequence 324, Application US/10177293
 ; Publication No. US20030124128A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lillie, James
 ; APPLICANT: Glatt, Karen
 ; APPLICANT: Zhao, Xumei
 ; APPLICANT: Gannavarpu, Manjula
 ; APPLICANT: Kamackar, Shubhangi
 ; APPLICANT: Mertens, Maureen
 ; APPLICANT: Myer, Vic
 ; APPLICANT: Wang, Youzhen
 ; APPLICANT: Xu, Yongyao
 ; APPLICANT: Hoersch, Sebastian
 ; APPLICANT: Monahan, John
 ; APPLICANT: Meyers, Rachel E.
 ; APPLICANT: Baer Jr., Robert C.
 ; APPLICANT: Horobegyl, Gabriel N.
 ; APPLICANT: Puzestay, Lajos
 ; APPLICANT: Meric, Aysegul
 ; APPLICANT: Sahin, Fudegul
 ; APPLICANT: Mills, Gordon B.
 ; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
 ; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
 ; FILE REFERENCE: MRI-038
 ; CURRENT APPLICATION NUMBER: US/10/177,293
 ; CURRENT FILING DATE: 2002-06-21
 ; PRIOR APPLICATION NUMBER: US 60/299,887

; PRIOR FILING DATE: 2001-06-21
 ; PRIOR APPLICATION NUMBER: US 60/301,572
 ; PRIOR FILING DATE: 2001-06-27
 ; PRIOR APPLICATION NUMBER: US 60/306,501
 ; PRIOR FILING DATE: 2001-07-18
 ; PRIOR APPLICATION NUMBER: US 60/325,002
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: US 60/362,585
 ; PRIOR FILING DATE: 2002-03-05
 ; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
 ; PRIOR FILING DATE: 2002-05-14
 ; NUMBER OF SEQ ID NOS: 506
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 324
 ; LENGTH: 394
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-177-293-324

Query Match 99.8%; Score 2067; DB 15; Length 394;
 Best Local Similarity 99.7%; Pred. No. 1,66-198;
 Matches 393; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSREMODVLAELVKRLVEKGETITGLLOEPDVEQODIETLHGSVHTLCTGPKGNRPVIL 60
 DB 1 MSREMODVLAELVKRLVEKGETITGLLOEPDVEQODIETLHGSVHTLCTGPKGNRPVIL 60
 QY 61 TYHDIGMNHKTCYNPLRYEEDMOEITQHFAVCHVDAPGQDGAASFPAGYMPSPMDOLAE 120
 DB 61 TYHDIGMNHKTCYNPLRYEEDMOEITQHFAVCHVDAPGQDGAASFPAGYMPSPMDOLAE 120
 QY 121 MLPGLVLOQFGLKSIIGMGTGAGAYILTRFALNPEMEVGLVLIIVNPCAEGMMDWAASKI 180
 DB 121 MLPGLVLOQFGLKSIIGMGTGAGAYILTRFALNPEMEVGLVLIIVNPCAEGMMDWAASKI 180
 QY 181 SGWTOALPDVNVSHLFGKEEMQSNVEVHTYRQHIIVDMNPGNLHLFINAVNSRRDLEIE 240
 DB 181 SGWTOALPDVNVSHLFGKEEMQSNVEVHTYRQHIIVDMNPGNLHLFINAVNSRRDLEIE 240
 QY 241 RPEMGTHTVTLQCPALLVVGDSPPAVDAVVECNKSLDPTKTTLLKMDCCGLPOLISQPAK 300
 DB 241 RPEMGTHTVTLQCPALLVVGDSPPAVDAVVECNKSLDPTKTTLLKMDCCGLPOLISQPAK 300
 QY 301 LAEAKFYVQGMGMPASMTRLMRSRTASGSSVTSLDGTRSRSHTSGTRSRSHTSGT 360
 DB 301 LAEAKFYVQGMGMPASMTRLMRSRTASGSSVTSLDGTRSRSHTSGTRSRSHTSGT 360
 QY 361 RSRHTSEGAHLDTTPNSGAAGNSAGPKSMEVSC 394
 DB 361 RSRHTSEGAHLDTTPNSGAAGNSAGPKSMEVSC 394

RESULT 6
 ; Sequence 156, Application US/10205194
 ; Publication No. US20030134301A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Warner-Lambert Company
 ; APPLICANT: Lee, Kevin
 ; APPLICANT: Dixon, Alistair
 ; APPLICANT: Brooksbank, Robert
 ; APPLICANT: Pinnock, Robert
 ; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
 ; FILE REFERENCE: WL-A-018201
 ; CURRENT APPLICATION NUMBER: US/10/205,194
 ; CURRENT FILING DATE: 5200-07-24
 ; PRIOR APPLICATION NUMBER: GB 0118354.0
 ; PRIOR FILING DATE: 2001-07-27
 ; NUMBER OF SEQ ID NOS: 177
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 156
 ; LENGTH: 394
 ; TYPE: PRT

ORGANISM: Mus musculus
 FEATURE:
 OTHER INFORMATION: Cytoplasmic protein Ndr1
 US-10-205-194-156

Query Match 94.6%; Score 1960; DB 12; Length 394;
 Best Local Similarity 93.7%; Pred. No. 9e-188;
 Matches 369; Conservative 14; Mismatches 11; Indels 0; Gaps 0;

QY 1 MSREMODVDLAELVKEKETTITGLQEPDVOEODITLHGSVHTLCGTPKGNRPVIL 60
 DB 1 MSREHVDVLAELVKEKETTITGLQEPDVOEODITLHGSVHTLCGTPKGNRPVIL 60
 QY 61 TYHDIGMNHKTCYNPLFNYEDMOEITQHFAVCHVADAPGQDGAASFPAGYMPSMDQLAE 120
 DB 61 TYHDIGMNHKTCYNPLFNYEDMOEITQHFAVCHVADAPGQDGAASFPAGYMPSMDQLAE 120
 QY 121 MLPGVLOOFGLSKSIIGMGTGAGAYILTRFALNPNPEMVEGLVILNPNPCABGMMDMAASKI 180
 DB 121 MLPGVLOOFGLSKSIIGMGTGAGAYILTRFALNPNPEMVEGLVILNPNPCABGMMDMAASKI 180
 QY 181 SGMTOALPDMVVSHLFGKEEMQSNVEVHTYRQHIIVDMNPGNLHLFTINAYNSRRDLEIE 240
 DB 181 SGMTOALPDMVVSHLFGKEEMQSNVEVHTYRQHIIVDMNPGNLHLFTINAYNSRRDLEIE 240
 QY 241 RPEMGTHTVTLQCPALLVVDSSPAVDVAVECNSKLDPTKTTLLKMACCGGLPOISQPAK 300
 DB 241 RPEMGTHTVTLQCPALLVVDSSPAVDVAVECNSKLDPTKTTLLKMACCGGLPOISQPAK 300
 QY 301 LAEAFKTYVQGMGYMPSASMTRLMRSRTASGSVTSIDGTRSRSHSSEGT 360
 DB 301 LAEAFKTYVQGMGYMPSASMTRLMRSRTASGSVTSIDGTRSRSHSSEGT 360
 QY 361 RSRSHSEGAHLDITPNSGAAGNSAGPKSMEVSC 394
 DB 361 RSRSHSEGAHLDITPNSGAAGNSAGPKSMEVSC 394

RESULT 7

US-10-108-260A-3803
 ; Sequence 3803, Application US/10108260A
 ; Publication No. US20040005560A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HELIX RESEARCH INSTITUTE
 ; TITLE OF INVENTION: NO. US20040005560A1 full length cDNA
 ; FILE REFERENCE: HI-A0106
 ; CURRENT APPLICATION NUMBER: US/10/108, 260A
 ; CURRENT FILING DATE: 2002-03-27
 ; NUMBER OF SEQ ID NOS: 5458
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3803
 ; LENGTH: 356
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-108-260A-3803

Query Match 87.7%; Score 1817; DB 12; Length 356;
 Best Local Similarity 90.1%; Pred. No. 1.7e-173;
 Matches 355; Conservative 0; Mismatches 1; Indels 38; Gaps 1;

QY 1 MSREMODVDLAELVKEKETTITGLQEPDVOEODITLHGSVHTLCGTPKGNRPVIL 60
 DB 1 MSREMODVDLAELVKEKETTITGLQEPDVOEODITLHGSVHTLCGTPKGNRPVIL 60
 QY 61 TYHDIGMNHKTCYNPLFNYEDMOEITQHFAVCHVADAPGQDGAASFPAGYMPSMDQLAE 120
 DB 61 TYHDIGMNHKTCYNPLFNYEDMOEITQHFAVCHVADAPGQDGAASFPAGYMPSMDQLAE 120
 QY 121 MLPGVLOOFGLSKSIIGMGTGAGAYILTRFALNPNPEMVEGLVILNPNPCABGMMDMAASKI 180
 DB 121 MLPGVLOOFGLSKSIIGMGTGAGAYILTRFALNPNPEMVEGLVILNPNPCABGMMDMAASKI 180
 QY 181 SGMTOALPDMVVSHLFGKEEMQSNVEVHTYRQHIIVDMNPGNLHLFTINAYNSRRDLEIE 240

DB 143 SGMTOALPDMVVSHLFGKEEMQSNVEVHTYRQHIIVDMNPGNLHLFTINAYNSRRDLEIE 202
 QY 241 RPEMGTHTVTLQCPALLVVDSSPAVDVAVECNSKLDPTKTTLLKMACCGGLPOISQPAK 300
 DB 203 RPEMGTHTVTLQCPALLVVDSSPAVDVAVECNSKLDPTKTTLLKMACCGGLPOISQPAK 262
 QY 301 LAEAFKTYVQGMGYMPSASMTRLMRSRTASGSVTSIDGTRSRSHSSEGT 360
 DB 263 LAEAFKTYVQGMGYMPSASMTRLMRSRTASGSVTSIDGTRSRSHSSEGT 322
 QY 361 RSRSHSEGAHLDITPNSGAAGNSAGPKSMEVSC 394
 DB 323 RSRSHSEGAHLDITPNSGAAGNSAGPKSMEVSC 356

RESULT 8

US-10-108-260A-3274
 ; Sequence 3274, Application US/10108260A
 ; Publication No. US20040005560A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HELIX RESEARCH INSTITUTE
 ; TITLE OF INVENTION: NO. US20040005560A1 full length cDNA
 ; FILE REFERENCE: HI-A0106
 ; CURRENT APPLICATION NUMBER: US/10/108, 260A
 ; CURRENT FILING DATE: 2002-03-27
 ; NUMBER OF SEQ ID NOS: 5458
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3274
 ; LENGTH: 323
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-108-260A-3274

Query Match 79.5%; Score 1646.5; DB 12; Length 323;
 Best Local Similarity 82.0%; Pred. No. 1.9e-156;
 Matches 323; Conservative 0; Mismatches 0; Indels 71; Gaps 1;

QY 1 MSREMODVDLAELVKEKETTITGLQEPDVOEODITLHGSVHTLCGTPKGNRPVIL 60
 DB 1 MSREMODVDLAELVKEKETTITGLQEPDVOEODITLHGSVHTLCGTPKGNRPVIL 60
 QY 61 TYHDIGMNHKTCYNPLFNYEDMOEITQHFAVCHVADAPGQDGAASFPAGYMPSMDQLAE 120
 DB 61 TYHDIGMNHKTCYNPLFNYEDMOEITQHFAVCHVADAPGQDGAASFPAGYMPSMDQLAE 120
 QY 121 MLPGVLOOFGLSKSIIGMGTGAGAYILTRFALNPNPEMVEGLVILNPNPCABGMMDMAASKI 180
 DB 121 MLPGVLOOFGLSKSIIGMGTGAGAYILTRFALNPNPEMVEGLVILNPNPCABGMMDMAASKI 180
 QY 181 SGMTOALPDMVVSHLFGKEEMQSNVEVHTYRQHIIVDMNPGNLHLFTINAYNSRRDLEIE 240
 DB 181 SGMTOALPDMVVSHLFGKEEMQSNVEVHTYRQHIIVDMNPGNLHLFTINAYNSRRDLEIE 240
 QY 241 RPEMGTHTVTLQCPALLVVDSSPAVDVAVECNSKLDPTKTTLLKMACCGGLPOISQPAK 300
 DB 241 RPEMGTHTVTLQCPALLVVDSSPAVDVAVECNSKLDPTKTTLLKMACCGGLPOISQPAK 300
 QY 301 LAEAFKTYVQGMGYMPSASMTRLMRSRTASGSVTSIDGTRSRSHSSEGT 360
 DB 301 LAEAFKTYVQGMGYMPSASMTRLMRSRTASGSVTSIDGTRSRSHSSEGT 360
 QY 361 RSRSHSEGAHLDITPNSGAAGNSAGPKSMEVSC 394
 DB 323 RSRSHSEGAHLDITPNSGAAGNSAGPKSMEVSC 323

RESULT 9

US-10-094-749-1842
 ; Sequence 1842, Application US/10094749
 ; Publication No. US20030219741A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ISOGAL, TAKAO

APPLICANT: SUGIYAMA, TOMOYASU
 APPLICANT: OTSUKI, TETSUJI
 APPLICANT: MAKAMATSU, AI
 APPLICANT: SATO, HIROYUKI
 APPLICANT: ISHII, SHIZUKO
 APPLICANT: YAMAMOTO, JUN-ICHI
 APPLICANT: ISONO, YUUKO
 APPLICANT: HIO, YURI
 APPLICANT: OTSUKA, KAORU
 APPLICANT: NAGAI, KEIICHI
 APPLICANT: IRIE, RYOTARO
 APPLICANT: TAMECHIKA, ICHIRO
 APPLICANT: SEKI, NAOHICO
 APPLICANT: YOSHIKAWA, TSUTOMU
 APPLICANT: OTSUKA, MOTOMYUKI
 APPLICANT: NAGAHARI, KENJI
 APPLICANT: MASUHO, YASUHIKO
 TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
 FILE REFERENCE: 08435/0160
 CURRENT APPLICATION NUMBER: US/10/094,749
 CURRENT FILING DATE: 2002-03-12
 PRIOR APPLICATION NUMBER: 60/350,435
 PRIOR FILING DATE: 2002-01-24
 PRIOR APPLICATION NUMBER: JP 2001-328361
 PRIOR FILING DATE: 2001-09-14
 NUMBER OF SEQ ID NOS: 3381
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 1842
 LENGTH: 339
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-094-749-1842

Query Match 54.9%; Score 1137.5; DB 12; Length 339;
 Best Local Similarity 59.6%; Pred. No. 2.5e-105;
 Matches 215; Conservative 50; Mismatches 67; Indels 29; Gaps 2;
 Db 34 EODIETLHGSVHTLCGTPEKGNRPVILTYHDIQNMHKTQCNPLFYVEDMOEITOHFAVCH 93
 8 EHDIEETPGGLHVVIRGSPKGNRPAILTYHDVGLNHLCEFTTFNFEDMOEITKHFFVCH 67
 Qy 94 VDAPOQDDAASFPAGYMPSPMDQLEMLPGVLOQFGLSIIGMGTGAGAYILTRFALNN 153
 Db 68 VDAPQOYGAQOFPGYQFPSEMEQLAANLPSVVGHFGRKYVIGIGVGAGAYLAKFALIF 127
 Qy 154 PEMVEGLVLINVPCEAGMMDMASKISGWTQALPDMVSHLFGKEEMQSNVEVHTYRQ 213
 Db 128 PDLVEGLVLINVIDPFGKGMIDMAATKLSGLTSLPDTVLSHLSQSEELVNNTELVOYSRQ 187
 Qy 214 HIVDMMNGNLHLFINANSRDLIERPMPGTHVTITLOCPALLVVGDSPPAVDAVVECN 273
 Db 188 QIGNVNANQQLFVNMVNSRRDLINRPVPAKTLRCFVMLVVGDNAPBEDGVVECN 247
 Qy 274 SKLPTKTLTKMADCGGLPOISQPAKLAFAFYVQGMGYMPASMTRLRSRTASGSS 333
 Db 248 SKLPTTTTFLKMDSGGLPQVTOGKLTLEAFKYLQGMGYMPASMTRLRSRTASLTS 307
 Qy 334 VTSLDGTRSRHTSEGRSRHTSEGRSRHTSEGRSGALHLDITTPNSGAAGSAGPKSMEVS 393
 Db 308 ASSVGSHPQACT-----HSESEGLGQVHNT-----MVS 338
 Qy 394 C 394
 Db 339 C 339

RESULT 10
 US-09-925-301-1404
 Sequence 1404, Application US/09925301
 Patent No. US2002032308A1
 GENERAL INFORMATION:
 APPLICANT: Rosen et al.
 TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

FILE REFERENCE: PA106
 CURRENT APPLICATION NUMBER: US/09/925,301
 CURRENT FILING DATE: 2001-08-10
 PRIOR APPLICATION NUMBER: PCT/US00/05882
 PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/124,270
 PRIOR FILING DATE: 1999-03-12
 NUMBER OF SEQ ID NOS: 1694
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 1404
 LENGTH: 251
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (37)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (41)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 US-09-925-301-1404

Query Match 53.2%; Score 1102; DB 9; Length 251;
 Best Local Similarity 97.7%; Pred. No. 5.6e-102;
 Matches 209; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Qy 103 AASFPAGYMPSPMDQLEMLPGVLOQFGLSIIGMGTGAGAYILTRFALNNPEMVEGLVL 162
 Db 35 AXLPRXYMPSPMDQLEMLPGVLOQFGLSIIGMGTGAGAYILTRFALNNPEMVEGLVL 94
 Qy 163 INVAPCAEGMMDMAASKISGWTQALPDMVSHLFGKEEMQSNVEVHTYRQHIIVNDMNG 222
 Db 95 INVAPCAEGMMDMAASKISGWTQALPDMVSHLFGKEEMQSNVEVHTYRQHIIVNDMNG 154
 Qy 223 NLHLFINANSRDLIERPMPGTHVTITLOCPALLVVGDSPPAVDAVVECNKLDPTKTT 282
 Db 155 NLHLFINANSRDLIERPMPGTHVTITLOCPALLVVGDSPPAVDAVVECNKLDPTKTT 214
 Qy 283 LKXADCGGLPOISQPAKLAFAFYVQGMGYMP 316
 Db 215 LKXADCGGLPOISQPAKLAFAFYVQGMGYMP 248

RESULT 11
 US-10-172-069-1
 Sequence 1, Application US/10172069
 Publication No. US20030167480A1
 GENERAL INFORMATION:
 APPLICANT: Stuart, Susan G.
 APPLICANT: Au-Young, Janice
 APPLICANT: Hillman, Jennifer L.
 APPLICANT: Shah, Puryi
 APPLICANT: Yue, Henry
 TITLE OF INVENTION: NDR2-RELATED PROTEINS
 FILE REFERENCE: PC-0038 CIP
 CURRENT APPLICATION NUMBER: US/10/172,069
 CURRENT FILING DATE: 2002-06-13
 PRIOR APPLICATION NUMBER: US/09/812,484
 PRIOR FILING DATE: 2001-03-19
 NUMBER OF SEQ ID NOS: 33
 SOFTWARE: PERL Program
 SEQ ID NO 1
 LENGTH: 371
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc feature
 OTHER INFORMATION: Incyte ID No. US20030167480A1 2227688CD1
 US-10-172-069-1

Query Match 50.6%; Score 1049; DB 12; Length 371;
 Best Local Similarity 52.1%; Pred. No. 2.1e-96;
 Matches 208; Conservative 64; Mismatches 89; Indels 38; Gaps 6;

QY 4 EMODVLAELVPLV-----EKGETITGLQEPDVQEDIEITLHGSVHTLGGTPKGN 55
 Db 3 ELQEVQITREKRLPBGQTPBEAKAEELARILLD-QGQTHSVETPYGVTFTVGTPEK 61
 QY 56 RPVILTYHIDGMNKTCTNPLFNEVDMOETIOHPAVCHVADPGQDGAASFPAGYMP 115
 Db 62 RPALITVHDVGLNYSKCFQPLFQEPDMOEIIQNEFRAVHVDAPGMEGAPVPLGYQYPSL 121
 QY 116 DQLEMLPGVLOQFGLKSIIGMGTGAGAYILTRFALNPNVEVGLVILNPNCAEGMDM 175
 Db 122 DQLDAMTICVQLYNFSTIIIGVGAGAYILARVALNHPDVEGLVILNIDPNKGMWDM 181
 QY 176 AASKISGWTQALPDMVNVSHLFGKEEMQSNVEVHTYRQHIYVNDMPGNLH.FINAYNSRR 235
 Db 182 AAHLTGLTSSIPENILGHLFSQELSGNSSELIQKYRNIITHAPLNDIEIYNSYNRR 241
 QY 236 DLEIRPMGTHVTTLQCPALLVVGDSAPVAVVECNKLDPTKTTLLKMA DCGGLPQI 295
 Db 242 DLNFER---GGDITLRCPVMLVVGDPAPHEDAVVECNKLDPTQTSFLKMA DSGGQPOL 297
 QY 296 SOPAKLAFAFYVQGMYPASMTRLMSRTASGSSVTSLDGTRSRSHTSEGT RSRSH 355
 Db 298 TOPGLTEAFKYPFQGMGYMASSCWTRLSRRTASLTSAASVDGNRSRSH----- 346
 QY 356 TSEGTSRSHTSEGAHLDTIPNSGAGNAGPKSMENVC 394
 Db 347 -----RTLSQSSSEG-----TLSSGPPGH-----TMEVSC 371

RESULT 12
 US-10-172-069-2
 ; Sequence 2, Application US/10172069
 ; Publication No. US20030167480A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Stuart, Susan G.
 ; APPLICANT: Au-Young, Janice
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Shah, Purvi
 ; APPLICANT: Yue, Henry
 ; TITLE OF INVENTION: NDR2-RELATED PROTEINS
 ; FILE REFERENCE: PC-0038 CIP
 ; CURRENT APPLICATION NUMBER: US/10/172, 069
 ; PRIOR FILING DATE: 2002-06-13
 ; PRIOR APPLICATION NUMBER: US/09/812, 484
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 2
 ; LENGTH: 402
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; OTHER INFORMATION: incycle ID No. US20030167480A1 3507515CD1
 ; NAME/KEY: unsure
 ; LOCATION: 328, 374, 379
 ; OTHER INFORMATION: unknown or other
 ; US-10-172-069-2

Query Match 50.1%; Score 1039; DB 12; Length 402;
 Best Local Similarity 55.1%; Pred. No. 2,4e-95;
 Matches 201; Conservative 59; Mismatches 85; Indels 20; Gaps 5;
 QY 4 EMODVLAELVPLV-----EKGETITGLQEPDVQEDIEITLHGSVHTLGGTPKGNRPVILTYH 63
 Db 3 ELQEVQITREKRLPBGQTPBEAKAEELARILLD-QGQTHSVETPYGVTFTVGTPEKPAIILTYH 55
 QY 56 RPVILTYHIDGMNKTCTNPLFNEVDMOETIOHPAVCHVADPGQDGAASFPAGYMP 123
 Db 62 RPALITVHDVGLNYSKCFQPLFQEPDMOEIIQNEFRAVHVDAPGMEGAPVPLGYQYPSLQDLMDIP 115

QY 124 GVLQOFLKSIIGMGTGAGAYILTRFALNPNVEVGLVILNPNCAEGMDMAASKISGW 183
 Db 116 CVIQLYNFSTIIIGVGAGAYILARVALNHPDVEGLVILNIDPNKGMWDMAAHLTGL 175
 QY 184 TQALPDMVNVSHLFGKEEMQSNVEVHTYRQHIYVNDMPGNLH.FINAYNSRRDIET RPPM 243
 Db 176 TSSIPENILGHLFSQELSGNSSELIQKYRNIITHAPLNDIEIYNSYNRRDLNFER-- 233
 QY 244 PGHTVTTLQCPALLVVGDSAPVAVVECNKLDPTKTTLLKMA DCGGLPQISQAPKLA 303
 Db 234 --GGDITLRCPVMLVVGDPAPHEDAVVECNKLDPTQTSFLKMA DSGGQPOLTOPGLT 291
 QY 304 AFKYPVQGMYPASMTRLMSRTASGSSVTSLDGTRSRSHT-SEGT RSRSH-----RS 354
 Db 292 AFKYPVQGMYPASMTRLMSRTASLTSAASVDGNRSRSH----- 351
 QY 355 HTSBG 359
 Db 352 HTWGG 356

RESULT 13
 US-10-172-069-32
 ; Sequence 32, Application US/10172069
 ; Publication No. US20030167480A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Stuart, Susan G.
 ; APPLICANT: Au-Young, Janice
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Shah, Purvi
 ; APPLICANT: Yue, Henry
 ; TITLE OF INVENTION: NDR2-RELATED PROTEINS
 ; FILE REFERENCE: PC-0038 CIP
 ; CURRENT APPLICATION NUMBER: US/10/172, 069
 ; PRIOR FILING DATE: 2002-06-13
 ; PRIOR APPLICATION NUMBER: US/09/812, 484
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 32
 ; LENGTH: 371
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; OTHER INFORMATION: Genbank ID No. US20030167480A1 96141566
 ; US-10-172-069-32

Query Match 50.1%; Score 1038; DB 12; Length 371;
 Best Local Similarity 52.1%; Pred. No. 2,7e-95;
 Matches 208; Conservative 64; Mismatches 89; Indels 38; Gaps 7;
 QY 4 EMODVLAELVPLV-----EKGETITGLQEPDVQEDIEITLHGSVHTLGGTPKGN 55
 Db 3 ELQEVQITREKRLPBGQTPBEAKAEELARILLD-QGQTHSVETPYGVTFTVGTPEK 61
 QY 56 RPVILTYHIDGMNKTCTNPLFNEVDMOETIOHPAVCHVADPGQDGAASFPAGYMP 115
 Db 62 RPALITVHDVGLNYSKCFQPLFQEPDMOEIIQNEFRAVHVDAPGMEGAPVPLGYQYPSL 121
 QY 116 DQLEMLPGVLOQFGLKSIIGMGTGAGAYILTRFALNPNVEVGLVILNPNCAEGMDM 175
 Db 122 DQLDAMTICVQLYNFSTIIIGVGAGAYILARVALNHPDVEGLVILNIDPNKGMWDM 181
 QY 176 AASKISGWTQALPDMVNVSHLFGKEEMQSNVEVHTYRQHIYVNDMPGNLH.FINAYNSRR 235
 Db 182 AAHLTGLTSSIPDMILGHLFSQELSGNSSELIQKYRNIITHAPLNDIEIYNSYNRR 241
 QY 236 DLEIRPMGTHVTTLQCPALLVVGDSAPVAVVECNKLDPTKTTLLKMA DCGGLPQI 295
 Db 242 DLNFER---GGDITLRCPVMLVVGDPAPHEDAVVECNKLDPTQTSFLKMA DSGGQPOL 297
 QY 296 SOPAKLAFAFYVQGMYPASMTRLMSRTASGSSVTSLDGTRSRSHTSEGT RSRSH 355

[illegible]

```

RESULT 14
US-10-104-047-2056
: Sequence 2056, Application US/10104047
: Publication No. US20030236392A1
: GENERAL INFORMATION:
: APPLICANT: HELIX RESEARCH INSTITUTE
: TITLE OF INVENTION: NO. US20030236392A1 full length cDNA
: FILE REFERENCE: H1-A0105
: CURRENT APPLICATION NUMBER: US/10/104,047
: CURRENT FILING DATE: 2002-03-25
: PRIOR APPLICATION NUMBER:
: PRIOR FILING DATE:
: NUMBER OF SEQ ID NOS: 4096
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 2056
: LENGTH: 311
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-10-104-047-2056

```

Query Match	43.1%;	Score 892.5;	DB 12;	Length 311;
Best Local Similarity	55.0%;	Pred. No. 8.3e-81;		
Matches 176;	Conservative 48;	Mismatches 67;	Indels 29;	Gaps 4

[illegible]

RESULT 15
US-10-264-237--2050
Sequence 2050, Application US/10264237
Publication No. US20040009491A1
GENERAL INFORMATION:
APPLICANT: Birse et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PAl31d1
CURRENT APPLICATION NUMBER: US/10/264,237
CURRENT FILING DATE: 2002-10-04
PRIOR APPLICATION NUMBER: PCT/US01/16450
PRIOR FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: US 60/205,515
PRIOR FILING DATE: 2000-05-19
NUMBER OF SEQ ID NOS: 2876
SOFTWARE: PatentIn Ver. 3.1

```

: SEQ ID NO 2050
: LENGTH: 325
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: MISC FEATURE
: LOCATION: (285)
: OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
:
: SS-10-264-237-2050
:

```

Query Match	41.3%	Score 855;	DB 12;	Length 325;
Best Local Similarity	-51.9%;	Pred. No. 5.1e-77;		
Matches 163;	Conservative 59;	Mismatches 74;	Indels 18;	Gaps 5

```

QY      4  EMOVDVLAELVKELVKGESITIGLLOEFVPOBODILETNGSFWHLVLCSTPKGNRVAIIITYH  63
Db      3  ELQEOVITEKEPLL-PCQTPEA-----AKTHSEVTPRGVSTFVYVCGTPKPKRRPAIIITYH  55

QY      64  DIGMNHKTQVNLFLPNYEMOEIOTHFVAVCHVDAQODGAASPAGMYTSPMOILAEMP  123
Db      56  DVGLANTSCQPLFQFQEDMOEIIQNFVRVADVADGEMEGAVFPLGQYRSLDDLAIMP  115

QY      124  GVLQOFGKSIIGMGTGAGAYILTRFALNPBMEYGLVLINVBCEAGMMDMAASKISGW  183
Db      116  CVLQYINFTIIIGVGVGAGAYILARVALNHPDVEGLVLINIDBNAGMMDMAHKLTGL  175

QY      184  TQALPNMVVSHFGKEEQMSNEVYHTRGQIVVNDMPGNHLIFINYNRRDLETFRPM  243
Db      176  TSSIFPMILIGHFSQEBELSGNSELQKRYRNIITRAPLNDLIELWNSYNNRRDLEFR--  233

QY      244  PGTHTVTLQCPALLVGDSSPAVDVAVBCNSKLDPTKITLLKMACCGLPQISQPAKLAE  303
Db      234  --GSDITTLRCPVMVLVVGDAQPHEDAVVEBCNSKLDPTQTSFLKMAKDSGGQPOLTX-ARQAD  290

QY      304  AFKTYVQAGMCMPS  317
Db      291  -----RGLQVLP  298

```

Search completed: January 29, 2004, 23:35:00
Job time : 41 secs


```

Query Match      16.7% Score 347; DB 2; length 325;
Best Local Similarity 30.4%; Pred. No. 7.9e-21;
Matches 101; Conservative 55; Mismatches 142; Indels 34; Gaps 11.

QY   27 LQEFVQOEIDITLHSGSVHTLCGTPKNGRRVILITYHDIGMNHKTCPNLFNEEDMQEIT 86
Db    6 LQVVVVQAQNCGLVHHVYQGNN--EEKGKTIILLVDHIGNHKS-FVRVNHPSMAIVK 62
      :|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY   87 QHFVAVCHDAFPGQDGAASFPAAGTGYRBSMDLAEMDGVTLQQEGLSIITGMGTGAGAYTL 146
      :|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

RESULT 5
D83263
probable hydrolytic enzyme PA3053 [imported] - *Pseudomonas aeruginosa* (strain PA01)
C/Species: *Pseudomonas aeruginosa*
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: D83263
 R:Stover, C.K.; Plam, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,
 ; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: D83263
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-335 <STO>
 A:Cross-references: GB:AE004730; GB:AE004091; NID:99949154; PIDN:AA06441.1; GSPDB:GN001
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA2953

Query Match 7.2%; Score 149; DB 2; Length 335;
 Best Local Similarity 25.1%; Pred. No. 0.00017;
 Matches 63; Conservative 43; Mismatches 85; Indels 60; Gaps 14;

Oy 104 ASPFAGWYPMDDLAEMLPGLQFGKLSIGGTGAGAYILTRPALNPEMVEGLVI 163
 Db 107 SSKPAHFOY-SFOQLAANTHALERLGVARSVIGSHSGMLATRYALLVPRQERLV- 164
 Oy 164 NVNFCAGGMDMAASKISGWTQALPDWVSHLFGKEKQSNVEVHTYRQ--HIVDMNP 221
 Db 165 -VNII--GLEDMKALGVF-W-RSYVD-----WYRRDQTSABGRQVQATYVAGEWRP 213
 Oy 222 GNHLFLIN-----AYNSRRDLIERPMPGHTTV-TLQCPALLVVG----- 260
 Db 214 -EPDRWYQMGWYRGKRGESVANNSALTVDIMFTQPVVYELDLQMPFTLLISEKONTA 272
 Oy 261 ---DSSPAVDVAVGNSKLDP-----TKTLLKMAADGGGPOISQPAKLAFAF 305
 Db 273 ICKDAAPA-----EUKAKLVYAOQLGKDAARRIPQATLVEFPDGHHTIOAPERFHOA- 326
 Oy 306 KYFVQGMGYMP 316
 Db 327 --LLEGLOTGP 335

RESULT 6
 H83276
 Pseudomonas aeruginosa (strain PA01)
 C:Species: *Pseudomonas aeruginosa*
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: H83276
 R:Stover, C.K.; Plam, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,
 ; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: H83276
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-315 <STO>
 A:Cross-references: GB:AE004721; GB:AE004091; NID:99949041; PIDN:AA06337.1; GSPDB:GN001
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA2949

Query Match 6.8%; Score 140; DB 2; Length 315;
 Best Local Similarity 22.9%; Pred. No. 0.00083;
 Matches 78; Conservative 49; Mismatches 116; Indels 98; Gaps 18;

Oy 10 LAEYVPLVEKGETTIGLQEPDVQODIET--LHGSVHTVLCGPPKGRPVILTYHDGM 67
 Db 27 LASVR-TYERG-LAG-LSHSVQVDNLEIAYLB-----GSK-NPTLLIHGGA 73
 Oy 68 ---NHKTCYNPLFNYEDMOBITTOHPAVCHVADPGQDGAAPFAGYMPMSDOLAEMLP 124
 Db 74 DKDWMLRFARPL-----TERHVVALDLPGGDSKPOQASY--DIVGTQAEVAVN 121

Oy 125 VLQDFGLKSIIGGTGAGAYILTRPALNPEMVEGLVIN---VNPFCAGGMDMAASKIS 181
 Db 122 PAALAGVRRLHLAGNSGSHIALYARHPQVLSLALINAGWMPARK----- 170
 Oy 182 GWTQALPDWVSHLFGKEKQSNVEVHTYRQ-----HIVDMNPGLHFLI-- 228
 Db 171 -----SELFEDLERGENPLVY---RQEDFPQKLDPFVFOQPPPLPAPLKRVLGE 216
 Oy 229 ---NAYNS-----RDLIERPMPGHTTVTLQCPALLVVGSSPAVDVAVGNSK 276
 Db 217 RAVASAFNNAQIFEQLRQRYTIPLEPPLP-----KTEAPTLILWGDRLDRLVLDV-----SSI 266
 Oy 277 DPTKTLTK-----MADCGLPQISQPAKLAFAKYPVQGM 312
 Db 267 EWRPRLKPSVIMENGVHWPVERBETRNQHTQATLDGV 307

RESULT 7 BVECBH

biotin biosynthesis protein biob - *Escherichia coli* (strain K-12)
 C:Species: *Escherichia coli*
 C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 01-Mar-2002
 C:Accession: J00081; G65136; S06071
 R:O'Regan, M.; Gloeckler, R.; Bernard, S.; Ledoux, C.; Ohnawa, I.; Lemoine, Y.
 Nucleic Acids Res. 17, 8004, 1989
 A:Title: Nucleotide sequence of the biob gene of *Escherichia coli*.
 A:Reference number: J00081; MUID:90016899; PMID:2678009
 A:Accession: J00081
 A:Molecule type: DNA
 A:Residues: 1-256 <ORF>
 A:Cross-references: GB:X15587; NID:941067; PIDN:CAA33612.1; PID:941068
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
 A.; Rose, D.T.; Mau, B.; Sha, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of *Escherichia coli* K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: G65136
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-256 <BLAT>
 A:Cross-references: GB:AE000417; GB:U00096; NID:92367220; PIDN:AACT6437.1; PID:G1789817;
 A:Experimental source: strain K-12, substrain MG1655
 C:Comment: This protein is involved in an early unknown step of the biotin biosynthesis
 C:Genetics:
 A:Gene: biob
 A:Map position: 75 min
 A:Superfamily: biob protein
 C:Keywords: biotin biosynthesis

Query Match 6.7%; Score 138.5; DB 1; Length 256;
 Best Local Similarity 24.2%; Pred. No. 0.00082;
 Matches 69; Conservative 35; Mismatches 80; Indels 101; Gaps 14;

Oy 42 GSVHTVLCGPPKGRPVILTYHDGMNKT--CYNPLFNYEDMOBITTOHPAVCHVADPG- 98
 Db 12 GNWHLV-----LHGWLNAEVRICD-----ELSSHFTLHVLDPGF 50
 Oy 99 ---QODGAAPFAGYMPMSDOLAEMLPGLQFGKLSIGGTGAGAYILTRPALNPE 155
 Db 51 GRSRGFGLS-----LADMAEAVLQCAPDKA-IWLGWSLGLVLSQALTHPE 97
 Oy 156 MYEGVLVILNVPFCAGGMDMAASK--ISGHTQALPDWVSHLFGKEKQSNVEVHTYR 212
 Db 98 RVQALVTVASSPCPSABDEWPGIRPDVLAGFQOQLSD-----DFQTVB----- 141
 Oy 213 OHIVDMNPGLHFLINAYNSRRDLIERPMPGHTTVTLQCPAL--LVGDSSPAVDVAV 270
 Db 142 -----RFLAQ--TMGTETARQDARALKYVLAALPMEVAVL- 176
 Oy 271 ECKSKLDPKTLTKMAADCGGLPQISQPAKLAFAKYPVQGMGM 315
 Db 177 --NGGLEILKTVDLRQ-----LQNVSMF-----FLRLYGYL 206

RESULT 8

AB1655

prolyl aminopeptidases homolog ltn1782 [imported] - *Listeria innocua* (strain Clip11262)C/Species: *Listeria innocua*

C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 06-Jan-2003

C/Accession: AB1655

R/Glasser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Jones, L.M.; Karsch, U.

Science 294, 849-852, 2001

A/Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schueller, T.; Simoes, N.; Tixeret, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A./Title: Comparative genomics of *Listeria* species.

A/Reference number: AB1077; MUID:21537279; PMID:11679669

A/Accession: AB1655

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-275 <GLA>

A/Cross-references: GB:AB005673; NID:g13423950; PIDN:AAK24382.1; GSPDB:GN00148

A/Experimental source: strain Clip11262

C/Genetics:

A/Gene: ltn1782

C/Superfamily: tropinesterase

Query Match 6.7%; Score 138.5; DB 2; Length 275; Best Local Similarity 21.4%; Pred. No. 0.00091; Matches 63; Conservative 49; Mismatches 124; Indels 59; Gaps 11;

QY 45 HTTLCGTPKGNRPVILTYHDIGMNHKTCYNPLEFNEY-MOEITQHFVCHVDAPGQODGA 103
 DB 7 HHHLNDISGEKFPVLMHGFQSSSET-----FQDISILKEHFSIIARDLGH--GK 57

QY 104 ASFPAGYWPMSDQLAEMLPVLOQFGKLSIIGMGTGAGAYILTRFALNPEWVEGLVLI 163
 DB 58 TSCPEKVARYSIENICDILASILHQLKIDSCFVLGYSWGVATFAAKYKELVGLVLI 117

QY 164 NNPCAEQMD--WASKISGWTQALPMVYVSHLFGKEMQSNVVEVHTYRQHLYNDMP 221
 DB 118 SSSP--GLRDEKAPASRISA-DNRLADLTLDN-----DGLRFPVAYWE----- 156

QY 222 GNLMFLINAYN---SRRLDEIERPMPGTHVT-----LQCPAL 257
 DB 157 -NLALFASQKLPALKKRIRLERLAQNPHGLAKSLRGWGTGKPSYENLADFTFPVLI 215

QY 258 VVGDSPPAVDAVVCNSKLDPTKTTLLKMDQCGLPQISQPAKLAFAKRYVQGM 312
 DB 216 ITGNIDKEFEKIAREMQLLP-NSTHYTVVQEGHVAVLEQPNIFSSQLIYMLEGI 269

RESULT 9

B87548

hypothetical protein CC2411 [imported] - *Caulobacter crescentus*C/Species: *Caulobacter crescentus*

C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C/Accession: B87548

R/Nierman, W.C.; Fiedlilyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.B.; Laud, M.T.; DeBoy, R.T.; Dodson, R.U.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolorn, J.; Brimicombe, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.; Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A/Title: Complete genome sequence of *Caulobacter crescentus*.

A/Reference number: AB7249; MUID:21173698; PMID:11259647

A/Accession: B87548

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-393 <STO>

A/Cross-references: GB:AB005673; NID:g13423950; PIDN:AAK24382.1; GSPDB:GN00148

C/Genetics:

A/Gene: CC2411

Query Match 6.5%; Score 135.5; DB 2; Length 393; Best Local Similarity 22.8%; Pred. No. 0.0027;

Matches 62; Conservative 47; Mismatches 112; Indels 51; Gaps 12;

QY 55 NRPVILTYHDIGMNHKTCYNPLEFNEYDMQITQHFVCHVDAPGQODGASFPAGYWP 114
 DB 20 DKPLVLNLSIGCD-LSLHDPV-----TPILTPDFRLRLIDTRGH--GASDAPSGDY--S 69

QY 115 MDQLAEMLPVLOQFGKLSIIGMGTGAGAYILTRFALNPEWVEGLVLIWVPCAEQMD 174
 DB 70 LDILADVLVMDAGAAKKTIGTSLGMIAMALASRAPDRVAVLACTSPAMDSD--- 126

QY 175 WASKISGWTQ-----ALPMVSHLFGKEMQSNVVEVHTYRQHLYNDMPN 223
 DB 127 -----SSWECORLAVIRAEGLSAIVEAWMSRFFSDDFRALHPEVETVRAGML-AQNPDG 179

QY 224 LHLPFNAYNRRLDEIERPMPGTHVTLLQCPALLVGDSPPA-----VDAVVCNSKLD 277
 DB 180 ---YCGCGAARMDALDRLP---KIAPTVLVTSKQVATPPFGHADRIY---AAMP 228

QY 278 PTKTTLKMDCCGGLPQISQPAKLAFAKRYFV 309
 DB 229 GARNAVIEAH---LPSLEAPAFAGAVRGFL 257

RESULT 10

F91160

biotin biosynthesis protein BioH [imported] - *Escherichia coli* (strain O157:H7, substrainC/Species: *Escherichia coli*

C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001

C/Accession: F91160

R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

Nasawa, N. 11-22, 2001

A/Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genom

A/Reference number: A99629; MUID:21156231; PMID:11258796

A/Accession: F91160

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-256 <HAY>

A/Cross-references: GB:BA000007; PIDN:BAH37677.1; PID:g13363728; GSPDB:GN00154

A/Experimental source: strain O157:H7, substrain RIMD 0509952

C/Genetics:

A/Gene: EC84255

C/Superfamily: bioH protein

Query Match 6.5%; Score 134.5; DB 2; Length 256; Best Local Similarity 23.9%; Pred. No. 0.0017; Matches 68; Conservative 36; Mismatches 80; Indels 101; Gaps 14;

QY 42 GSVHVTLCGTPKGNRPVILTYHDIGMNHKTCYNPLEFNEYDMQITQHFVCHVDAPG- 98
 DB 12 GNHVLV-----LHMGJLNAEYVRICD-----ELSSHFTLLHVDLPGF 50

QY 99 ---QODGASFPAGYWPMSDQLAEMLPVLOQFGKLSIIGMGTGAGAYILTRFALNPE 155
 DB 51 GRSRGFALS-----LAEMAEVALRQAPDKA-TLWMSIGGLVASQIALTHPE 97

QY 156 MVEGLVLIWVPCAEQMDAAK---ISGWTQALPMVSHLFGKEMQSNVVEVHTYR 212
 DB 98 RVQALVTVASPPCSARDEWPGIKPVDLAGFQQLSD-----DFQRYE----- 141

QY 213 QHIVNDMPGNLHFLINAYNRRLDEIERPMPGTHVTLLQCPAL--LVVGDSSPAVDVAV 270
 DB 142 -----RFLALQ--TWGTEIARQDARLAKKTVLALPMEVDVL- 176

RESULT 11

E86006

biotin biosynthesis protein BioH [imported] - *Escherichia coli* (strain O157:H7, substrainC/Species: *Escherichia coli*

C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 27-Nov-2001
 C:Accession: E86006
 R:Perma, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew
 Iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A>Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: E86006
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-256 <STO>
 A:Cross-references: GB:AE005174; MID:g12518042; PIDN:AAG58513.1; GSPDB:GN00145; UWGP:247
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: bioH
 C:Superfamily: bioH protein

Query Match 6.5%; Score 134.5; DB 2; Length 256;
 Best Local Similarity 23.9%; Pred. No. 0.0017;
 Matches 68; Conservative 36; Mismatches 80; Indels 101; Gaps 14;

QY 42 GSVHTLCTGPKGNRPVLTTHDGMNKT--CYNPLRYEEDMOETTHFAVCHVDAG- 98
 DB 12 GNVHLVLT-----LHGMGLNABVWRCID-----EELSHFTLHVLDTLDFG 50
 QY 99 ---QODGAAPGAGYVPSMDQLAEMLPVLOQFGLKSLIGMGAGAYILTRPALNPE 155
 DB 51 GRSRFGALIS-----LAMAARVLRQADPKA-ILKMSLGLVLSQILTPTE 97
 QY 156 MVEGLVLINVPCEAGMDMAASK--ISGWTQALPDVNVSHLFKEEMQSNVEVHTYR 212
 DB 98 RVQALVTVAASPCFARDEMPKIDVLAGFQQLSD-----DFQRIVE----- 141
 QY 213 QHIVNDMPGNLHLFINAVNSRDLERPEMPGHTTTLQCPAL--LVYGDSSPAVDVAV 270
 DB 142 -----RFLALQ--TWGTTARQDARALKKTVLALPMPEDVL- 176
 QY 271 ECNSKLDPTKTTLLKMAADCGGLPOLISOPAKLAFAKVFVQGGVM 315
 DB 177 --NGLEILKTYDLRLP-----LQNVPM-----FLRLVGYL 206

RESULT 12
 AB1284
 prolyl aminopeptidases homolog lmo1674 (imported) - listeria monocytogenes (strain EGD-e)
 C:Species: listeria monocytogenes
 C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 06-Jan-2003
 C:Accession: AB1284
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bioecker
 ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
 Science 294, 849-852, 2001
 A:Authors: Kretz, T.; Kuhn, M.; Kunat, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
 ok, C.; Schlueter, J.; Simeos, N.; Tietzer, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
 A>Title: Comparative genomes of listeria species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AB1284
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-275 <GLA>
 A:Cross-references: GB:NC_003210; PIDN:CAC99752.1; PID:g16411110; GSPDB:GN00177
 A:Experimental source: strain EGD-e
 C:Genetics:
 A:Gene: lmo1674
 C:Superfamily: tropinesterase

Query Match 6.2%; Score 128.5; DB 2; Length 275;
 Best Local Similarity 20.3%; Pred. No. 0.0061;
 Matches 59; Conservative 48; Mismatches 129; Indels 55; Gaps 11;

QY 45 HTTLCGTPKGNRPVLTTHDGMNKTCTNPLFNYD--MOETTHFAVCHVDAPGQODGA 103
 DB 7 HVHLLTAISGEKRALMLHGFTGTSET-----FQDSISGLKGRFNIIAPDLGH--GN 57

QY 104 ASPGAGYVPSMDQLAEMLPVLOQFGLKSLIGMGAGAYILTRPALNPEVEGLVLI 163
 DB 58 TASBEELSSYTMENICEDLAGILHLQLVNSRCPVLYGMSGAVARAFATPFKRRLGILV 117
 QY 164 NVNCEAGMDMAASKISGWTQALPDVNVSHLFKEEMQSNVEVHTYRQHI VDNMNGN 223
 DB 118 SSSPGLF--QEDIRASRVEA--DNRLADWI-----ELEGIVPRVDYWE-----N 157
 QY 224 LHLF-----INAVNSRDLERPEMPGHTTTL-----QC-----PALIV 259
 DB 158 LALPASQKVSPEMKRIRSRSLQNSHGLMSLRGKGTGKQPSYNNCLANFTPEVLLIT 217
 QY 260 GDSSPAVDAYECSKLDPTKTTLLKMAADCGGLPOLISOPAKLAFAKVFVQ 310
 DB 218 GALDEKFEKIAQEMHQLLP--NSTVSIQEGHAYVLEQPPNSFSSQLNWMLE 267

RESULT 13
 CB2692
 biotin biosynthesis protein xpf356 (imported) - xylella fastidiosa (strain 9a5c)
 C:Species: xylella fastidiosa
 C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
 C:Accession: CB2692
 R:anonymous, The xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
 Nature 406, 151-157, 2000
 A>Title: The genome sequence of the plant pathogen xylella fastidiosa.
 A:Reference number: A82515; MUID:20365717; PMID:10910347
 A>Note: for a complete list of authors see reference number A59328 below
 A:Accession: CB2692
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-255 <STM>
 A:Cross-references: GB:AE003967; GB:AE003849; MID:g9106347; PIDN:AAE84165.1; GSPDB:GN001
 A:Experimental source: strain 9a5c
 R:Simson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
 Briones, M.R.S.; Bueno, M.R.P.; Canarigo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre, H
 as-Neto, E.; Docena, C.; El-Dorriy, H.; Facinanci, A.P.; Ferreira, A.J.S.
 submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Petro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Frohm
 U.D.; Junqueira, M.L.; Kemper, E.L.; Kitzajima, J.P.; Krieger, J.B.; Kuramae, E.E.; Laig
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.,
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.B.; de Sa, R.G.; Santelli, R.V.; Sawasak
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
 M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
 A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: xpf356
 C:Superfamily: bioH protein

Query Match 6.2%; Score 127.5; DB 2; Length 255;
 Best Local Similarity 22.2%; Pred. No. 0.0066;
 Matches 64; Conservative 43; Mismatches 106; Indels 75; Gaps 15;

QY 54 GNRPVILTTHDGMNKTCTNPLFNYEDMOETTHFAVCHVDAPGQODGAAPGAGYV 113
 DB 7 GYCPALVLIHGMAN-HSGVFAPL-----VEQLRAHHTLVLVDLDEGH-----GYVHT 51
 QY 114 SMDQALEMLPGVLOQFGLKSLIGMGAGAYILTRPALNPEVNGVLYVINVPCLEG 171
 DB 52 TLTPLA--LPRVVAIAAATPPAVWLMSLGLFALNAATLPQ--VGLILMLAATPPCVR 108
 QY 172 WMDNA-ASKISGWTQALPDVNVSHLFKEEMQSNVEVHTYRQHI VDNMNGN 220
 DB 109 REDMPHAYVSIFFQ-----FAEDLKONTERTINRLADLTGISTAGSELR--- 155
 QY 221 PGNLHLFINAVNSRDLERPEMPGT-----HT-----VTLCALLVGDSS 263
 DB 156 --QROILNA-----NHTENTATLQAGLELALHTDLRRRAVIDLTLPSPSLWIAQORD 203

Oy 264 PAVDAV-VECNKSLDPT-KTTLKMGADCGGLPQISQPAKLAFAKRYFV 309
 Db 204 RLVPAASIAHATAPASGQTELLTITGGHAPFLSHANQMTALQHF1 251

RESULT 14

AF0016
 probable biotin biosynthesis protein bioH (imported) - Yersinia pestis (strain CO92)
 C/Species: Yersinia pestis
 C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
 C/Accession: AF0016
 R/Parikhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
 deno-Tarrag, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G.;
 11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,
 Nature 413, 523-527, 2001
 A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A/Reference number: AB0001; MUID:21470413; PMID:11586360
 A/Accession: AF0016
 A/Status: Preliminary
 A/Molecule type: DNA
 A/Residues: 1-258 <KUR>
 A/Cross-references: GB:AL590842; PIDN:CAC88992.1; PID:G15978234; GSPDB:GN00175
 C/Genetics:
 A/Gene: bioH
 C/Superfamily: bioH protein

Query Match 6.0%; Score 125; DB 2; Length 258;
 Best Local Similarity 23.1%; Pred. No. 0.011; Mismatches 112; Indels 64; Gaps 16;
 Matches 66; Conservative 44;

Oy 49 CGTPKGRPVILTYHDIGMNHKT--CYNPLFNYEDMOEITQHFAVCHVDAPG---QDDG 102
 Db 9 CG--EGDCDLVL-LHGMGLNAGVWHCI-----IDRLAPHFRLHLDLPYGRSQDYG 57
 Oy 103 AASFPAGMTYPSMDQLAEMLPGLKSLIGGTGAGAYILTRFALNNPEWEGVLV 162
 Db 58 AMS-----LADMAERVAQ--APKQALWLGWSMGLVVASQIALSQPECRCGLIT 104
 Oy 163 INVNPACAGMMDMAASK---ISGWTQALPD-----MVVSHLFGKEMQSNVYVHT- 210
 Db 105 VSSSPCFARDEWPCIKREVLAFGPHQSLDDPHRTVERFLAQTLGTSSRQDARLKSV 164
 Oy 211 YRQHIIVDMN--PGNLHFLINAYNSRRDLEIERMPGTHVTTLQCPALLVVGDSPPAVDA 268
 Db 165 VLQHQMPDVEVLTGLAIL-----RTADLRITLAG-----FTLPFRVYGH---LDS 208
 Oy 269 VV--ECNKLDPPT--KTTLKMGADCGGLPQISQPAKLAFAKRYFVQ 310
 Db 209 LVPRKVASLLDSAMPQTSVVMGAAHAAPFISHPDPFAKLINFAE 254

RESULT 15

AE0097
 probable biotin biosynthesis protein STY4287 (imported) - Salmonella enterica subsp. ent
 C/Species: Salmonella enterica subsp. enterica serovar Typh
 A/Note: this species has also been called Salmonella typhi
 C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C/Accession: AE0097
 R/Parikhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher,
 th, T.; Connerly, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
 A/Reference number: AB0502; MUID:21534947; PMID:11677608
 A/Accession: AE0097
 A/Status: Preliminary
 A/Molecule type: DNA
 A/Residues: 1-256 <PAR>
 A/Cross-references: GB:AL513382; PIDN:CAD08105.1; PID:G16505084; GSPDB:GN00176
 C/Genetics:
 A/Gene: STY4287
 C/Superfamily: bioH protein

Query Match 5.9%; Score 123; DB 2; Length 256;
 Best Local Similarity 23.3%; Pred. No. 0.015;
 Matches 67; Conservative 35; Mismatches 100; Indels 86; Gaps 14;

Oy 53 KGRNPVILTYHDIGMNHKT--CYNPLFNYEDMOEITQHFAVCHVDAPGQDDGAASPPAGY 110
 Db 11 EGNCHLVL-LHGMGLNAGVWHCI-----EELGSHFRLHLDLPYGRSSGFGAMT 60
 Oy 111 MYPSMDQLAEMLPGLKSLIGGTGAGAYILTRFALNNPEWEGVLINVPCC-- 168
 Db 61 LBEXTAQVAKNAP-----DQALWLGWSLGLVAGMALTPHERVQALVTASSPCFS 112
 Oy 169 -ABGMDMAASKISGWTQALPDVNVSHLFGKEMQSNV-----VHATYRQHIIVDM 219
 Db 113 ARBWPQIKREIIGFPQQLSD-----DPQRTVERFLAQTLGTETARQDA----- 158
 Oy 220 NPGNLHFLINAYNSRRDLEIERMPGTH-----TVTLQ-----CPALLVVGDS 262
 Db 159 -----RTLSKVLAQMPDVEVLNGLLEILKTVDLRBALKNVMPFLRLYG-- 204
 Oy 263 SPAVDAYECNKLDPPTKTTLK-----MADCGGLPQISQPAKLAFA 304
 Db 205 -YLDGLVP--RKIAPLDLTPHSTSQIMAKAHAAPFISHPAFCA 248

Search completed: January 29, 2004, 23:29:51
 Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 23:25:08 ; Search time 17 Seconds
(without alignments)
1089.913 Million cell updates/sec

Title: US-10-057-832-2
Perfect score: 2072
Sequence: 1 MSREMQVDVLAEVKPLVEKG.....TPNSGAAGNAGPKMEVSC 394

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2072	100.0	394	1 NR1_HUMAN	Q92597 homo sapien
2	1960	94.6	394	1 NR1_MOUSE	Q62433 mus musculu
3	1254	60.5	375	1 NR3_HUMAN	Q9UGV2 homo sapien
4	1244	60.0	375	1 NR3_MOUSE	Q9UGY9 mus musculu
5	1117	53.9	352	1 NR4_HUMAN	Q9A1P0 homo sapien
6	1108	53.5	352	1 NR4_RAT	Q9Z219 rattus norv
7	1049	50.6	371	1 NR2_HUMAN	Q9AUN36 homo sapien
8	1038	50.1	371	1 NR2_MOUSE	Q9UGY90 mus musculu
9	351.5	17.0	352	1 SR21_HELAN	O23969 helianthus
10	347	16.7	325	1 VPJ1_CABEL	O02485 caenorhabdi
11	138.5	6.7	256	1 B10H_ECOLI	P13001 escherichia
12	107.5	5.2	1808	1 TENA_CHICK	P10039 gallus galli
13	99.5	4.8	331	1 ILVQ_SYNEU	Q86G90 synechococc
14	99.5	4.8	543	1 MIG1_KLUMA	P52288 kluyveromyc
15	99	4.8	523	1 DHE4_CHLMO	P28998 chlorella s
16	98	4.7	966	1 CAPP_MEDSA	O02735 medicago sa
17	97.5	4.7	1353	1 VGL2_CYHOC	P36334 human coron
18	95	4.6	1159	1 RPOC_FORCN	O33431 porphyromon
19	94.5	4.6	750	1 YKS7_YEAST	P34231 saccharomyc
20	94.5	4.6	1242	1 IRS1_HUMAN	P33568 homo sapien
21	93.5	4.5	294	1 DEH1_MORSP	O01398 moraxella s
22	93.5	4.5	530	1 INR2_BOVIN	O95141 bos taurus
23	93.5	4.5	575	1 HEMA_SEBUN	P04853 sendai viru
24	93.5	4.5	1486	1 MUXB_ECOLI	P22523 escherichia
25	92.5	4.4	1850	1 GLPT_HAEIN	P02845 gallus galli
26	92	4.4	1850	1 VIT2_CHICK	P06863 sendai viru
27	91.5	4.4	575	1 HEMA_SEBUN	O18663 dictyosteli
28	91.5	4.4	1321	1 IRFX_DICDI	P00632 actinobact
29	90.5	4.4	266	1 E1H2_MACTA	Q57427 haemophilus
30	90.5	4.4	287	1 Y193_HAEIN	O16623 homo sapien
31	90	4.3	238	1 SFR7_HUMAN	Q00238 rattus norv
32	89.5	4.3	545	1 ICA1_RAT	P27562 sendai viru
33	89.5	4.3	575	1 HEMA_SEBUN	

ALIGNMENTS

RESULT 1	ID	NR1_HUMAN	STANDARD;	PRT;	394 AA.
AC	Q92597	015207; Q9NYR6; Q9UK29;			
DT	15-JUL-1999	(Rel. 38, Created)			
DT	15-JUL-1999	(Rel. 38, Last sequence update)			
DT	15-SEP-2003	(Rel. 42, Last annotation update)			
DE	NRG1 protein (N-myc downstream regulated gene 1 protein)				
DE	(Differentiation-related gene 1 protein) (DRG1) (Reducing agents and tunicamycin-responsive protein) (RTP) (Nickel-specific induction protein Cap43) (Ric42).				
DE	NRG1 OR RTP OR DRG1 OR CAP43.				
GN	Homo sapiens (Human).				
OS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutazoa; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
NP	SEQUENCE FROM N.A.				
RC	TISSUE=Umbilical vein endothelial cells;				
RX	MEDLINE=97094664; PubMed=8939898;				
RA	Kokame K., Kato H., Miyata T.;				
RT	"Homocysteine-responsive genes in vascular endothelial cells identified by differential display analysis. GRP78/BiP and novel genes.";				
RT	J. Biol. Chem. 271:29659-29665(1996).				
RL	[2]				
RN	SEQUENCE FROM N.A.				
RP	MEDLINE=97395543; PubMed=9251681;				
RX	van Belzen N., Dinjens W.N.M., Drieseld M.P.G., Groen N.A.,				
RA	van der Made A.C.J., Nozawa Y., Vlietstra R., Trepman U.,				
RT	Boeman F.T.;				
RT	"A novel gene which is up-regulated during colon epithelial cell differentiation and down-regulated in colorectal neoplasms.";				
RT	Lab. Invest. 77:85-92(1997).				
RL	[3]				
RN	SEQUENCE FROM N.A.				
RP	TISSUE=Lung;				
RX	MEDLINE=98266801; PubMed=9605764;				
RA	Zhou D., Salnikow K., Costa M.;				
RT	"Cap43, a novel gene, specifically induced by Ni2+ compounds.";				
RT	Cancer Res. 58:2182-2189(1998).				
RL	[4]				
RN	SEQUENCE FROM N.A.				
RP	MEDLINE=99326165; PubMed=10395947;				
RX	Pignatelli D., Joulia D., Balaguer P., Basset A., Marti J., Comnes T.;				
RA	"Differential expression of the RTP/DRG1/NRG1 gene product in proliferating and growth arrested cells.";				
RT	Biochim. Biophys. Acta 1450:364-373(1999).				
RL	[5]				
RN	SEQUENCE FROM N.A.				
RP	TISSUE=Kidney;				
RX	MEDLINE=22388257; PubMed=12477932;				
RA	Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,				
RT	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,				
RT	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,				
RT	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heise F.,				

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stjepanec M., Soares M.B., Bonaldo M.F., Cacaevant T.L., Scheetz T.B.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carinini P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwen F.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hilaly S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallie D.E.,
 RA Schermer A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [6]
 RP SEQUENCE OF 1-198 FROM N.A.
 RA Blechschmidt K., Kalaydjieva L., Goodman R., Gresham D., Baas F.,
 RA Jonge R.D., Schilhabel M., Schattevoy R., Baumgart C., Menzel U.,
 RA Rosenthal A.;
 RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
 RN [7]
 RP SEQUENCE OF 1-21 FROM N.A.
 RC TISSUE=Brain;
 RA Angelicheva D., Kalaydjieva L.;
 RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
 RN [8]
 RP FUNCTION.
 RX MEDLINE=99438066; PubMed=9766676;
 RA Kudistean S.K., Arizti P., Reimer C.L., Sugrue M.M., Aaronson S.A.,
 RA Lee S.W.;
 RT "Inhibition of tumor cell growth by RTP/rict4 and its responsiveness
 to p53 and DNA damage.";
 RL Cancer Res. 58:4439-4444 (1998).
 CC -1- FUNCTION: MAY HAVE A GROWTH INHIBITORY ROLE.
 CC -1- SUBCELLULAR LOCATION: WHEREAS IN PROSTATE EPITHELIUM AND PLACENTAL
 CC CHORION IT IS LOCATED IN BOTH THE CYTOPLASM AND THE NUCLEUS,
 CC NUCLEAR STAINING IS NOT OBSERVED IN COLON EPITHELIUM CELLS.
 CC INSTEAD ITS LOCALIZATION CHANGES FROM THE CYTOPLASM TO THE PLASMA
 CC MEMBRANE DURING DIFFERENTIATION OF COLON CARCINOMA CELL LINES IN
 CC VITRO.
 CC -1- TISSUE SPECIFICITY: UBIQUITOUS; EXPRESSED MOST PROMINENTLY IN
 CC PLACENTAL MEMBRANES AND PROSTATE, KIDNEY, SMALL INTESTINE, AND
 CC OVARY TISSUES. REDUCED EXPRESSION IN ADENOCARCINOMAS COMPARED TO
 CC NORMAL TISSUES. IN COLON, PROSTATE AND PLACENTAL MEMBRANES, THE
 CC CELLS THAT BORDER THE LUMEN SHOW THE HIGHEST EXPRESSION.
 CC -1- INDUCTION: BY HOMOCYSTEINE, 2-MERCAPTOETHANOL, TUNICAMYCIN IN
 CC ENDOTHELIAL CELLS. INDUCED APPROXIMATELY 20-FOLD DURING IN VITRO
 CC DIFFERENTIATION OF THE COLON CARCINOMA CELL LINES HT29-D4 AND
 CC CACO-2. INDUCED BY NICKEL COMPOUNDS IN ALL TESTED CELL LINES. THE
 CC PRIMARY SIGNAL FOR ITS INDUCTION IS AN ELEVATION OF FREE
 CC INTRACELLULAR CALCIUM ION CAUSED BY NICKEL ION EXPOSURE. OKADAIC
 CC ACID, A SERINE/THREONINE PHOSPHATASE INHIBITOR, INDUCED ITS
 CC EXPRESSION MORE RAPIDLY AND MORE EFFICIENTLY THAN NICKEL.
 CC -1- SIMILARITY: BELONGS TO THE NDRG FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL: D87953; BAI13505.1; -;
 DR EMBL: X92845; CA63430.1; -;
 DR EMBL: AF004162; AAC13419.1; -;
 DR EMBL: AF186190; -; NOT ANNOTATED_CDS.
 DR EMBL: BC003175; AAH03175.1; -;
 DR EMBL: AF203080; AAF1305.1; -;
 DR GenBank: HGNC:7679; NDRG1.
 DR MIM: 605262; -;

DR GO: 0008181; F: tumor suppressor; TAS.
 DR GO: 0009412; P: response to heavy metal; TAS.
 DR InterPro: IPR004142; Ndr.
 DR InterPro: IPR003737; Ser_estr_site.
 DR Pfam: PF03096; Ndr; 1.
 KW Nuclear protein; Repeat.
 FT DOMAIN 339 368
 FT REPEAT 339 348
 FT REPEAT 349 358
 FT REPEAT 359 368
 FT CONFLICT 145 145
 FT SEQUENCE 394 AA; 42835 MW; 4C816B9C85E3756F CRC64;
 SQ
 Query Match 100.0%; Score 2072; DB 1; Length 394;
 Best Local Similarity 100.0%; Pred. No. 4.5e-162;
 Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSRRMDVDLAEVKPLVEKGETITGLQEPDVOEODIETLHGSVHTLCGTPKGNRPVIL 60
 DB 1 MSRRMDVDLAEVKPLVEKGETITGLQEPDVOEODIETLHGSVHTLCGTPKGNRPVIL 60
 QY 61 TYHDIGNNHKTCTNPLFNIEDMOEITQHPAVCHVDAPGQDGAASPAGTYRSMQALAE 120
 DB 61 TYHDIGNNHKTCTNPLFNIEDMOEITQHPAVCHVDAPGQDGAASPAGTYRSMQALAE 120
 QY 121 MLPGVLOQFGLKSLIGTGAGAYILTRPALNPNWEGVLINVPNCAEGMDMAASKI 180
 DB 121 MLPGVLOQFGLKSLIGTGAGAYILTRPALNPNWEGVLINVPNCAEGMDMAASKI 180
 QY 181 SGWTOALPDWVSHLPEKEMQSNVEVHTYRQHIYNDMNGMLHFINAYNSRRDLEIE 240
 DB 181 SGWTOALPDWVSHLPEKEMQSNVEVHTYRQHIYNDMNGMLHFINAYNSRRDLEIE 240
 QY 241 RPMPGHTTTLQCPALLVVDSSPAVDVVECKSKDPTTTLTKNADCGGLPQISPAK 300
 DB 241 RPMPGHTTTLQCPALLVVDSSPAVDVVECKSKDPTTTLTKNADCGGLPQISPAK 300
 QY 301 LAEAFKTFVQGMGMPASATRLMRSTRASGSSTSLDGTGRSHTSEGRSHTSEGT 360
 DB 301 LAEAFKTFVQGMGMPASATRLMRSTRASGSSTSLDGTGRSHTSEGRSHTSEGT 360
 QY 361 RSRSHSTSEGAHLDTTPNSGAAGNSAGPKSMVEVC 394
 DB 361 RSRSHSTSEGAHLDTTPNSGAAGNSAGPKSMVEVC 394
 RESULT 2
 NDRL_MOUSE
 ID NDRL_MOUSE STANDARD; PRT; 394 AA.
 AC 062433; P97862;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE NDRL1 protein (N-myc downstream regulated gene 1 protein) (Protein
 DE Ndrl1).
 GN NDRL1 OR NDRL1 OR TDS.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99310663; PubMed=10381566;
 RA Shimono A., Okuda T., Kondoh H.;
 RT "N-myc-dependent repression of ndrl, a gene identified by direct
 RT subtraction of whole mouse embryo cDNAs between wild type and N-myc
 RT mutant.";
 RL Mech. Dev. 83:39-52 (1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Hydridoma;
 RX MEDLINE=97289706; PubMed=9144177;

RA Lin T.-M., Chang C.;
 RT "Cloning and characterization of TDD5, an androgen target gene that is
 RT differentially repressed by testosterone and dihydrotestosterone.";
 RT Proc. Natl. Acad. Sci. U.S.A. 94:4988-4993(1997).
 CC -1- FUNCTION: MAY HAVE A GROWTH INHIBITORY ROLE (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (by similarity).
 CC -1- TISSUE SPECIFICITY: EXPRESSED MOST ABUNDANTLY IN KIDNEY FOLLOWED
 CC BY NERVOUS TISSUES (HYPOTHALAMUS, CEREBELLUM, AND CEREBRUM) AND
 CC PREPUTIAL GLAND, AND IN MUCH SMALLER QUANTITIES IN OTHER TISSUES,
 CC EXCEPT DUODENUM AND PROSTATE.
 CC -1- DEVELOPMENTAL STAGE: ITS EXPRESSION INCREASES AFTER 13.5 DPC WHEN
 CC N-MYC EXPRESSION IS LARGELY DOWN-REGULATED.
 CC -1- INDUCTION: REPRESSED BY TESTOSTERONE AND ALSO IN A LESSER EXTENT
 CC BY DIHYDROTESTOSTERONE.
 CC -1- SIMILARITY: BELONGS TO THE NDRG FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC
 CC EMBL; U60593; AAB03484.1; -;
 CC EMBL; U52073; AAB58249.1; -;
 CC MGD; MGI:1341799; Ndr1.
 CC InterPro; IPR004142; Ndr.
 CC Pfam; PF03096; Ndr; 1.
 CC Nuclear protein; Repeat.
 CC DOMAIN
 CC 339 368
 CC
 CC REPEAT 339 348
 CC REPEAT 349 358
 CC REPEAT 359 368
 CC
 CC CONFLICT 33 35 OEO -> LEE (IN REF. 2).
 CC CONFLICT 89 89 O -> P (IN REF. 2).
 CC CONFLICT 103 109 ASPFVG -> PLPSQM (IN REF. 2).
 CC CONFLICT 141 148 AGAVILTR -> PMKLHDP (IN REF. 2).
 CC CONFLICT 191 208 VVSHFGKEEINNVV -> CVPEPLXGDTQRCGM
 CC (IN REF. 2).
 CC
 CC CONFLICT 241 241 R -> A (IN REF. 2).
 CC CONFLICT 298 350 PAKLAEPFYVQGMGYPASMTLRMSRTASGSSVTSLE
 CC GTRSRSHSECP -> RSLRLPSSSLCRHGTHARQCQDPS
 CC DRVPHVMLOCHILGT (IN REF. 2).
 CC
 CC SEQUENCE 394 AA; 43008 MW; 905CA11EC4C67C2 CR64;
 CC
 CC Query Match 94.6%; Score 1960; DB 1; Length 394;
 CC Best Local Similarity 93.7%; Pred. No. 6-6e-153;
 CC Matches 369; Conservative 14; Mismatches 11; Indels 0; Gaps 0;

DB 301 LAEAFKYPVQGMGYPASMTLRMSRTASGSSVTSLEGTRSRSHSECPGPRSRHSTSGS 360
 QY 361 RSRSHSECAHLDITPNSGAAGNAGPKSMVSC 394
 DB 361 RSRSHSEDAHLDITPNSGAAGNAGPKSMVSC 394
 RESULT 3
 NDR3_HUMAN STANDARD; PRT; 375 AA.
 AC Q99GQ2; Q96P18; Q9BY77; Q9H3N7; Q9H411; Q9H8J6;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE NDRG3 protein.
 GN NDRG3
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND TISSUE SPECIFICITY.
 RX MEDLINE=21251006; PubMed=11352569;
 RA Zhou R.-H., Kokame K., Tsukamoto Y., Yutani C., Kato H., Miyata T.;
 RT "Characterization of the human NDRG gene family: a newly identified
 RT member, NDRG4, is specifically expressed in brain and heart.";
 RL Genomics 73:86-97(2001).
 RP [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RA Mao Y., Xie Y., Zhou Z., Zhao W., Zhao S., Wang W., Huang Y., Wang S.,
 RA Tang R., Chen X., Wu C.;
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Spleen;
 RX MEDLINE=21933870; PubMed=11936845;
 RA Qu X., Zhai Y., Wei H., Zhang C., Xing G., Yu Y., He F.;
 RT "Characterization and expression of three novel
 RT differentiation-related genes belong to the human NDRG gene family.";
 RL Mol. Cell. Biochem. 229:35-44(2002).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 3).
 RC TISSUE=Placenta;
 RA Iisagai T., Oca T., Hayashi K., Sugiyama T., Oeuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Wagatsuma M., Hosofiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuho Y.,
 RA Niomiya K., Iwayanagi T.;
 RT "NEDO human cDNA sequencing project";
 RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggaley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Bead L.M., Beare D.N.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clisg S., Cooley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Frazer A., French L., Garner P.,
 RA Grahame D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Lehaealsho M.H., Leverisha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McComachie L.J., McEay K., McMurtry A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,

RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
 RA Skuse C.D., Smith M.L., Soderlund C., Steward C.A., Sulterson J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Williams S.A.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.,
 RA "The DNA sequence and comparative analysis of human chromosome 20.",
 RL Nature 414:865-871(2001).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=3;
 CC Name=1;
 CC IsoId=O9UGV2-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=O9UGV2-2; Sequence=VSP_003419;
 CC Name=3;
 CC IsoId=O9UGV2-3; Sequence=VSP_003420;
 CC -1- TISSUE SPECIFICITY: Ubiquitous. Highly expressed in brain.
 CC -1- SIMILARITY: BELONGS TO THE NDRG FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB044943; BAB2067.1; -;
 DR EMBL; AF251054; AAK3494.1; -;
 DR EMBL; AF08609; AAL08807.1; -;
 DR EMBL; AK023618; BAB14620.1; -;
 DR EMBL; AL132768; CAC10098.2; -;
 DR EMBL; AL031662; CAB5625.1; ALT_SEQ.
 DR Genew; HGNC:14462; NDRG3.
 DR MIM; 605273; -;
 DR GO; GO:0005737; C:cytoplasm; IDA.
 DR GO; GO:0030154; P:cell differentiation; NAS.
 DR GO; GO:0030308; P:negative regulation of cell growth; NAS.
 DR GO; GO:0007283; P:spermatogenesis; NAS.
 DR InterPro; IPR004142; Ndr.
 DR InterPro; IPR000379; Ser_estrs_site.
 DR Pfam; PF03096; Ndr; 1.
 KM Alternative splicing.
 FT VARSPLIC 20 31 Missing (in isoform 2).
 FT VARSPLIC 47 135 Missing (in isoform 3).
 FT VARSPLIC 122 122 E -> G (IN REF. 3).
 FT CONFLICT 32 32 P -> S (IN REF. 2).
 FT CONFLICT 122 122 P -> S (IN REF. 2).
 SQ SEQUENCE 375 AA; 41408 MW; 237B8DF43AAB7F1E CRC64;
 Query Match 60.5%; Score 1254; DB 1; Length 375;
 Best Local Similarity 61.8%; Pred. No. 3.6e-95;
 Matches 244; Conservative 58; Mismatches 67; Indels 26; Gaps 6;
 QY 4 EMOVDVLAIVKPVLEKGETITGLQGFVDOEDIEHTLHGSVHTLTGTPKGNRPVLTTH 63
 DB 3 ELQDVQLTEIKPLL-NDKNGTRNPFDDCOEHDIETHGVHTIRGLPKGNRPVLTTH 61
 QY 64 DIGMNHKTCYCNPLFNVYEDMOEITQHPAVCHVAPGQDGAASFPAGYMPSPMDLAEMLP 123
 DB 62 DIGLNHKSCTFAFPNFEDMOEITQHPAVCHVAPGQDGAASFPPTGYQYPTMDLAEMLP 121
 QY 124 GYLQOFGKLSITGMGTAGAYITLTPALNPNPEVVEGLVINVPNCAEGMDMAASKISGW 183
 DB 122 PVLTHLSIKSIIGVGAGAYILSRFALNHEVEGLVILNVDPKAGWIDMAASKISGL 181
 QY 184 TQALPDVNVSHLFEKGEKQSNVVEVHTYRQHTIVDMMNGULHFLINVNSRRDLEIRPM 243
 DB 182 TTNVVDITLHNFHGFQBEIQANLDLITQYRMHTADINQDNLQFLINSYNGRDLLEIRPI 241
 QY 244 PGTH--TVTLQCPALLVVGDSPPAVDAVVECNKSLDPTXTLLTKMADCGGLPQISQPAK 300

DB 242 LQGNDRKSKTLKSTLLVVDGNSPAVEAVVECNRLNPIINTLLTKMADCGGLPQVVGPK 301
 QY 301 LAEFKFKVVGGMGMPASNTRLMRST-ASGSVSUJSDOTRRSRSHSEGTRRSRHSEG 359
 DB 302 LTFEKFLLGGMGYIPASWTRLARSRHSTSSLSGSESPFSRSVLT-----SNQSDG 354
 QY 360 TRSRSHSEGAHLDITPNSGAAGNSAGPKSMETSC 394
 DB 355 TQESCEP-----DVLDRH-----QTEVESC 375
 RESULT 4
 NDR3 MOUSE
 ID NDR3 MOUSE STANDARD; PRT; 375 AA.
 AC QGVYF;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE NDRG3 protein (Ndr3 protein).
 GN NDRG3 OR NDR3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=20050077; PubMed=10581191;
 RA Okuda T., Kondoh H.;
 RT "Identification of new genes ndr2 and ndr3 which are related to
 RT Nr1/RTF/Drg1 but show distinct tissue specificity and response to
 RT N-myc";
 RL Biochem. Biophys. Res. Commun. 266:208-215 (1999).
 CC -1- TISSUE SPECIFICITY: THYMUS, NERVOUS SYSTEM, GUT AND KIDNEY.
 CC -1- DEVELOPMENTAL STAGE: ITS EXPRESSION IS ALREADY SIGNIFICANT AT 9.5
 CC DPC, COVERING THE ENTIRE EMBRYO EXCEPT THE HEART, AND IT SHOWS
 CC ONLY A SLIGHT INCREASE IN LATER DEVELOPMENTAL STAGES.
 CC -1- SIMILARITY: BELONGS TO THE NDRG FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB033922; BAA5883.1; -;
 DR MGI; MGI:1352499; Ndr3.
 DR GO; GO:0005737; C:cytoplasm; ISS.
 DR GO; GO:0030154; P:cell differentiation; ISS.
 DR GO; GO:0030308; P:negative regulation of cell growth; ISS.
 DR GO; GO:0007283; P:spermatogenesis; ISS.
 DR InterPro; IPR004142; Ndr.
 DR InterPro; IPR000379; Ser_estrs_site.
 DR Pfam; PF03096; Ndr; 1.
 SQ SEQUENCE 375 AA; 41555 MW; B33D3CC3E816AEAI CRC64;
 Query Match 60.0%; Score 1244; DB 1; Length 375;
 Best Local Similarity 61.0%; Pred. No. 2.4e-94;
 Matches 241; Conservative 58; Mismatches 70; Indels 26; Gaps 6;
 QY 4 EMOVDVLAIVKPVLEKGETITGLQGFVDOEDIEHTLHGSVHTLTGTPKGNRPVLTTH 63
 DB 3 ELQDVQLTEIKPLL-NDKNGTRNPFDDCOEHDIETHGVHTIRGLPKGNRPVLTTH 61
 QY 64 DIGMNHKTCYCNPLFNVYEDMOEITQHPAVCHVAPGQDGAASFPAGYMPSPMDLAEMLP 123
 DB 62 DIGLNHKSCTFAFPNFEDMOEITQHPAVCHVAPGQDGAASFPPTGYQYPTMDLAEMLP 121
 QY 124 GYLQOFGKLSITGMGTAGAYITLTPALNPNPEVVEGLVINVPNCAEGMDMAASKISGW 183

DB 122 PVLTLMSKSIIGICVAGAVILSRFALNHPVLVGLVILNIDPCAKMIDMAKSLGCF 181
 QY 184 TQALPDVTVSHLFGKEENQSVNVEVTVYRHOHIVNDMPGNLHLPFNAYNSRRDLEIERPM 243
 DB 182 TTNVITDIILAHFGEELQANLNDLQYTRHIAQINQENLQPLFGSTNGRDLIEIERPI 241
 QY 244 PG---THTVTLQCPALLVGVSSPAVDVVECNSTKLDPTKTLTLMADCGGLPQISQPAK 300
 DB 242 LGQNDNRKLTLCSTLLVGVDSPAVEAVVECNSTRDPTINTTLLTMADCGGLPQVQPGK 301
 QY 301 LAEAFKYVQCMGVPASMTRLKMSRT-ASGSSVTSLDGTRSRKHTSEGRSRKHTSEG 359
 DB 302 LTEAKRYLQCGXGYPASMTRLASRTHTSSISGSGSPSPRSVT-----SNQSDG 354
 QY 360 TRSRHTSEGAHLDTTPNSGAAGNAGPKSMVSC 394
 DB 355 TQESCESP-----DVLDRH-----QTMVEVSC 375

RESULT 5
 NDRA_HUMAN STANDARD: PRT: 352 AA.
 AC Q9ULP0: Q96PL9: Q9GZM1: Q9GZN3: Q9GZX0;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE NDRG4 protein (Brain development-related molecule 1) (Vascular smooth
 muscle cell associated protein-8) (SNAP-8).
 GN NDRG4 OR BDM1 OR KIAA1180.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
 RX MEDLINE=21251006; Pubmed=11352569;
 RA Zhou R.-H., Kokame K., Tsukamoto Y., Yutani C., Kato H., Miyata T.;
 RT "Characterization of the human NDRG gene family: a newly identified
 member, NDRG4, is specifically expressed in brain and heart";
 RL Genomics 73:66-97(2001).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 3).
 RC TISSUE=Heart;
 RA Nishimoto S., Tawara J., Aoki T., Toyoda H., Komurasaki T.;
 RT "Molecular cloning and characterization of the human vascular smooth
 muscle cell associated protein-8 (SNAP-8).";
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 4).
 RC TISSUE=Spleen;
 RX MEDLINE=21933870; Pubmed=11936845;
 RA Qu X., Zhai Y., Wei H., Zhang C., Xing G., Yu Y., He F.;
 RT "Characterization and expression of three novel
 differentiation-related genes belong to the human NDRG gene family";
 RL Mol. Cell. Biochem. 229:35-44(2002).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Amygdala;
 RX MEDLINE=21154917; Pubmed=11230166;
 RA Wisman S., Weil B., Wellenreuther R., Gassenhuber J., Glasl S.,
 Ansoerge W., Boecker M., Bloecker H., Bauersachs S., Blum H.,
 Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
 Mewes H.-W., Ottenwaelder B., Obermaier A., Tampe J., Heubner D.,
 Wandt R., Korn B., Klein M., Poustka A.;
 RT "Towards a catalog of human genes and proteins: sequencing and
 analysis of 500 novel complete protein coding human cDNAs";
 RL Genome Res. 11:422-435(2001).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Brain;
 RX MEDLINE=22158633; Pubmed=12168954;
 RA Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.;
 RT "Construction of expression-ready cDNA clones for KIAA genes: manual

RT curation of 330 KIAA cDNA clones.";

RL DNA Res. 9:99-106(2002).
 RN [6]
 RP SEQUENCE OF 80-352 FROM N.A. (ISOFORM 1).
 RC TISSUE=Brain;
 RX MEDLINE=20039618; Pubmed=10574461;
 RA Hirotsawa M., Nagase T., Ishikawa K.-I., Kikuno R., Nomura N.,
 Ohara O.;
 RT "Characterization of cDNA clones selected by the GenMark analysis
 from size-fractionated cDNA libraries from human brain";
 RL DNA Res. 6:329-336(1999).
 RN [7]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Muscle;
 RX MEDLINE=22388257; Pubmed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 Brownstein M.J., Umed T.B., Toshiyuki S., Carninci P., Prange C.,
 Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullen S.J.,
 Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,
 Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 Fahey J., Heltan E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 Butlerfield Y.S.N., Krzywinski M.I., Skliska U., Smallos D.B.,
 Schercher A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: MAY PLAY A ROLE IN THE EARLY POSTNATAL DEVELOPMENT AND
 CC -!- FUNCTION OF NEURONAL CELLS (BY SIMILARITY).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=4;
 CC Name=1; Synonyms=NDRG4-Bvar;
 CC IsoId=G9ULP0-1; Sequence=Displayed;
 CC Name=2; Synonyms=NDRG4-B;
 CC IsoId=G9ULP0-2; Sequence=VSP_003422;
 CC Name=3; Synonyms=NDRG4-H;
 CC IsoId=G9ULP0-3; Sequence=VSP_003421, VSP_003422;
 CC Name=4;
 CC IsoId=G9ULP0-4; Sequence=VSP_003423;
 CC -!- TISSUE SPECIFICITY: Expressed only in brain and heart. Isoforms 1
 CC and 2 are only expressed in brain. Isoform 3 is expressed in both
 CC heart and brain.
 CC -!- SIMILARITY: BELONGS TO THE NDRG FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL; AB044947; BAB20071.1; -
 CC EMBL; AB044947; BAB20072.1; -
 CC EMBL; AB044947; BAB20073.1; -
 CC EMBL; AB044944; BAB20068.1; -
 CC EMBL; AB044944; BAB20069.1; -
 CC EMBL; AB044945; BAB20070.1; -
 CC EMBL; AB044946; BAB20071.1; -
 CC EMBL; AB021172; BAB20288.1; -
 CC EMBL; AF108608; AAL08806.1; -
 CC EMBL; AL136584; CAB66519.1; -
 CC EMBL; AB033006; BAB6494.2; ALT_INT.
 CC EMBL; BC011795; AAH11795.1; -
 CC Genew; HGNC:14466; NDRG4.
 CC GO; GO:0005737; C:cytoplasm; NAS.

```

DR GO; GO:0030154; P:cell differentiation; NAS.
DR GO; GO:0016049; P:cell growth; NAS.
DR GO; GO:0006950; P:response to stress; NAS.
DR InterPro; IPR004142; Ndr.
DR Pfam; PF03096; Ndr; 1.
KM Developmental protein; Alternative splicing.
FT DOMAIN 253
FT VARSPLIC 1 256
POLY-THR.
MPECDWG -> MAGLOELRPPPEKPLRGQDATELESSDAP
LAAADTWK (in isoform 3).
/FTid=VSP 003421.
FT VARSPLIC 289 302
IAYLKDRLESGAV -> M (in isoform 2 and
isoform 3).
/FTid=VSP 003422.
FT VARSPLIC 291 303
Missing (in isoform 4).
/FTid=VSP 003423.
SQ SEQUENCE 352 AA; 38459 MW; 4CFAFC7C820013B6 CRC64;

Query Match 53.9%; Score 1117; DB 1; Length 352;
Best Local Similarity 57.2%; Pred. No. 5.4e-84;
Matches 214; Conservative 51; Mismatches 67; Indels 42; Gaps 3;

QY 34 EODIETLHSGVHTLCGTGKGNRPVILTYHDIGMNHKTCYNPLFNYEDMOETQHFAVCH 93
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 8 EHDIEPTGLHVIKIRSGPKGNRPALITLTHDVGHNHLCEFTFPNFBDMOETKHFAVCH 67
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 94 VDAFGQDGAASFPAGYMPSPMDLAEMLPGLVQOGLKSIIGMGAGAYILTRFALNN 153
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 68 VDAFGQDGAASFPAGYMPSPMDLAEMLPGLVQOGLKSIIGMGAGAYILTRFALNN 127
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 154 PEMVEGLVILNVPKAGMMDMAASKISGWTQALPDMVSHLFGKEKMOSEVAVHTYRQ 213
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 128 PDVEGLVILNVIDPDKGMDMAATKLSGLTSTLPDVTLSHLSFQSEELVNNTELVSQYRQ 187
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 214 HIVNDMPGNLHLFTIANYNSRDLERPMGCHTTLTLOCPALLVGDSSPAVDAYVECN 273
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 188 QIGNVVQANLQLFWMNYSRRDLINPGTVPAKTLRCFVWLVDGNAFADGVVECN 247
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 274 SKLDPKTLTLKMAADCGGLPQISOPAKLAFAKIFYVQMGY-----MPSASM 320
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 248 SKLDPKTLTLKMAADCGGLPQVTPGKLTFAKIFYLQMGYIAHLKDRLSGAVPASM 307
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 321 TRLMRSTASGSSVTSIDGTRSRSHSEGTSRSHSEGTSRSHSEGAHLDTTPNSGA 380
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 308 TRLMRSTASLTSASVDSRPPACT-----HSESESGMGVNH----- 347
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 381 AGNSAGPKSMVEVC 394
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 348 -----MEVSC 352
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |

RESULT 6
NDR4 RAT STANDARD; PRT; 352 AA.
ID NDR4_RAT STANDARD; PRT; 352 AA.
AC Q922L9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NDRG4 protein (Brain development-related molecule 1).
GN NDRG4 OR BDM1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=9255674; PubMed=10320792;
RA Yamauchi Y., Hongo S., Nishinaka N., Ohashi T., Takahashi R.,
RA Takeda M.;
RT "Molecular cloning and characterization of a novel developmentally
regulated gene, Bdm1, showing predominant expression in postnatal rat
brain.";
RL Brain Res. Mol. Brain Res. 68:149-158(1999).

```

```

CC -1- FUNCTION: MAY PLAY A ROLE IN THE EARLY POSTNATAL DEVELOPMENT AND
CC FUNCTION OF NEURONAL CELLS.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE BRAIN AND HEART. WEAKLY IN
CC THE KIDNEY, MOST PROMINENTLY IN POSTNATAL BRAIN WHERE IT IS
CC EXPRESSED WIDELY IN THE OLFACTORY BULB, CEREBRAL CORTEX,
CC HIPPOCAMPUS, CEREBELLUM, THALAMUS, AND MEDULLA OBLONGATA.
CC -1- SIMILARITY: BELONGS TO THE NDRG FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AF045564; M202415.1; -.
DR InterPro; IPR004142; Ndr.
DR Pfam; PF03096; Ndr; 1.
KM Developmental protein.
FT DOMAIN 253
FT VARSPLIC 1 256
POLY-THR.
SQ SEQUENCE 352 AA; 38487 MW; EF3SD773B9D255D9 CRC64;

Query Match 53.5%; Score 1108; DB 1; Length 352;
Best Local Similarity 57.2%; Pred. No. 2.9e-83;
Matches 214; Conservative 50; Mismatches 68; Indels 42; Gaps 3;

QY 34 EODIETLHSGVHTLCGTGKGNRPVILTYHDIGMNHKTCYNPLFNYEDMOETQHFAVCH 93
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 8 EHDIEPTGLHVIKIRSGPKGNRPALITLTHDVGHNHLCEFTFPNFBDMOETKHFAVCH 67
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 94 VDAFGQDGAASFPAGYMPSPMDLAEMLPGLVQOGLKSIIGMGAGAYILTRFALNN 153
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 68 VDAFGQDGAASFPAGYMPSPMDLAEMLPGLVQOGLKSIIGMGAGAYILTRFALNN 127
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 154 PEMVEGLVILNVPKAGMMDMAASKISGWTQALPDMVSHLFGKEKMOSEVAVHTYRQ 213
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 128 PDVEGLVILNVIDPDKGMDMAATKLSGLTSTLPDVTLSHLSFQSEELVNNTELVSQYRQ 187
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 214 HIVNDMPGNLHLFTIANYNSRDLERPMGCHTTLTLOCPALLVGDSSPAVDAYVECN 273
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 188 QISSVNVQANLQLFWMNYSRRDLINPGTVPAKTLRCFVWLVDGNAFADGVVECN 247
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 274 SKLDPKTLTLKMAADCGGLPQISOPAKLAFAKIFYVQMGY-----MPSASM 320
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 248 SKLDPKTLTLKMAADCGGLPQVTPGKLTFAKIFYLQMGYIAHLKDRLSGAVPASM 307
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 321 TRLMRSTASGSSVTSIDGTRSRSHSEGTSRSHSEGTSRSHSEGAHLDTTPNSGA 380
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 308 TRLMRSTASLTSASVDSRPPACT-----HSDSESGMGVNH----- 347
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 381 AGNSAGPKSMVEVC 394
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 348 -----MEVSC 352
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |

RESULT 7
NDR2 HUMAN STANDARD; PRT; 371 AA.
ID NDR2_HUMAN STANDARD; PRT; 371 AA.
AC Q9UN36; Q96FD3; Q96FT0; Q96PNO; Q9B0H5; Q9ULH2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-FEB-2003 (Rel. 42, Last annotation update)
DE NDRG2 protein (Syd1709613 protein).
GN NDRG2 OR SYD1 OR KIAA1248.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Brain;
RA Deng Y.C., Yao L.B., Su C.Z., Lui X.P.;

```

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP TISSUE=Brain.
 RC MEDLINE=20039619; PubMed=10574462;
 RA Nagase T., Ishikawa K.-I., Kikuno R., Hiroseawa M., Nomura N.,
 RA Ohara O.
 RT "Prediction of the coding sequences of unidentified human genes. XV.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro."
 RL DNA Res. 6:337-345(1999).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Amygdala;
 RX MEDLINE=21154917; PubMed=11230166;
 RA Wieman S., Weil B., Wellenreuther R., Gassenhuber J., Glaesl S.,
 RA Ansoerg W., Boeher M., Bloecher H., Baerachis S., Blum H.,
 RA Lauber J., Duesenheoft A., Beyer A., Koehrer K., Strack N.,
 RA Meves H.-W., Ottenwelder B., Obermair B., Tampe J., Heubner D.,
 RA Mambrot R., Korn B., Klein M., Pousetka A.;
 RT "Towards a catalog of human genes and proteins: sequencing and
 RT analysis of 500 novel complete protein coding human cDNAs."
 RL Genome Res. 11:422-435(2001).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=121251006; PubMed=11355569;
 RA Zhou R.-H., Kokame K., Tsukamoto Y., Yutan C., Kato H., Miyata T.;
 RT "Characterization of the human NDRG gene family: a newly identified
 RT member, NDRG4, is specifically expressed in brain and heart."
 RL Mol. Cell. Biochem. 229:35-44(2002).
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Brain;
 RA Liu X., Deng Y., Yao L.;
 RT "Cloning and analysis of human genomic DNA sequence of N-myc
 RT downstream regulator 2,"
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 3).
 RC TISSUE=Brain;
 RX MEDLINE=22386257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Bluetow K.H., Scheafer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hoeh F.,
 RA Datchenko L., Mandina K., Farmer A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Tomihlyki S., Carninci P., Prange C.,
 RA Rana S.S., Loughlino N.A., Peters G.J., Abramson J.D., Mullany S.J.,
 RA Bosak S.A., McGwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Rahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Bakkeney R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Buttefield J.S.N., Krzywnski M.I., Skalska U., Smallue D.E.,
 RA Schuchet A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- EVENT=ALTERNATIVE PRODUCTS;
 CC Name=1;
 CC IsoId=09UNJ36-1; Sequence=Displayed;
 CC Name=2; Synonym=NDRG2var;

CC		Isoid=Q9UN36-2; Sequence=VSP_003417;
CC	Name=3;	
CC	IsoId=Q9UN36-3; Sequence=VSP_003418;	
CC	-1 SIMILARITY: BELONGS TO THE NDRG FAMILY.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.ebi-sib.ch/announce/	
CC	or send an email to license@ebi-sib.ch).	
CC	-----	
DR	EMBL; AF159092; AA043131.2; -	
DR	EMBL; AB033074; BAA86562.1; ALT_INIT.	
DR	EMBL; AL336574; CAB66509.1; -	
DR	EMBL; AF304051; AAL08624.1; -	
DR	EMBL; AY028430; AAK50340.1; -	
DR	EMBL; BC010458; AAH10458.1; -	
DR	EMBL; BC011240; AAH11240.1; -	
DR	Genew; HGNC:14460; NDRG2.	
DR	MIM; 605272; -	
DR	GO; GO:0005829; C:cytosol; NAS.	
DR	GO; GO:0030154; P:cell differentiation; ISS.	
DR	InterPro; IPR004142; Ndr.	
DR	Pfam; PF03096; Ndr; 1.	
KW	Alternative splicing.	
FT	VARSPLIC 26	Missing (in isoform 2). /FRTD=VSP_003417. /FRTID=VSP_003418.
FT	VARSPLIC 262 272	/FRTID=VSP_003418. /FRTID=VSP_003418.
FT	CONFLICT 48 48	G -> V (IN REF. 7).
FT	CONFLICT 54 54	V -> A (IN REF. 5).
FT	CONFLICT 172 172	D -> G (IN REF. 6).
FT	CONFLICT 250 250	D -> R (IN REF. 5).
FT	CONFLICT 296 296	O -> N (IN REF. 6).
FT	CONFLICT 306 306	A -> V (IN REF. 5).
SQ	SEQUENCE 371 AA; 40798 MW; 7B4950B12BD3F595 CRC64;	
Query Match	50.6%; Score 1049; DB 1; Length 371;	
Best Local Similarity	52.1%; Pred. No. 2,1e-78;	
Matches 208; Conservative 64; Mismatches 89; Indels 38; Gaps 6		
QY	4 EMQDVDAEKKPLV-----EKETITGLLOEPVQOEDITFIAGSVHTLTCTGPKN 55	
DB	3 ELGEVOITEKRPILLPGQTPEAKKAELAARILD-QGQTHSVERTPYSVTFTVGTPPK 61	
QY	56 RPVLITLTHDGNNHKTCYNPLFNVEDNQEIQHVAVCNDAPGGODGAASFPAGYTFSM 115	
DB	62 RPALLITHADVLNKKSCFOPLPQEDMQEIIQNVRHVHVAFGMEBGAVPEPLQYOYPSL 121	
QY	116 DOLAEMLPGLVQOGLKSIIIGMGAGAYILTRPALNPENVEGVYLINVPAEGMMD 175	
DB	122 DOLADMIPCVLYQNYNESTIIGVGAGAYILARIALNHPTVBGLVINIDPAKGMDW 181	
QY	176 AASKISGMTALPDMMVSHLGKEKMSNVENVHYTROHIVDNMPGNLIFINYSRR 235	
DB	182 AAHLTGTLTSIPMIILGHLSFGSELSGNSLLIKYNIITHANLDNIELYMNSYNNR 241	
QY	236 DLIERPMGTHTYTLOCPLLTVGDSSPAVDAVECNSKIDPTKTTLTKRADCGGLPOI 295	
DB	242 DLNER-----GSDITLRCPVNLVYGDOAPHEDAVECNSKIDPTQTFCLKADSGQPQL 297	
QY	296 SOPKLAEAFKYFQOGNGVPASMTMLMSRTTASSGYSLDSTRSRSHSBCTRSSH 355	
DB	298 TOPKLAEAFKYFQOGNGVPASMTMLMSRTTASSGYSLDSTRSRSHSBCTRSSH 346	
QY	356 TSEGRSRSHSBGAHLDTIPNSGAAGNASGPKMEVSC 394	
DB	347 -----RTLQSSESG-----TLSSPPGH-----TMVSC 371	

```

NDR2_MOUSE
ID_NDR2_MOUSE STANDARD; PRT; 371 AA.
AC_O9QYG0.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NDRG2 protein (Ndr2 protein).
DE NDRG2 OR NDR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=20050077; PubMed=10581191;
RA Okuda T.; Kondoh H.;
RT "Identification of new genes ndr2 and ndr3 which are related to
RT Ndr1/RTPr1 but show distinct tissue specificity and response to
RT N-myc."
RL Biochem. Biophys. Res. Commun. 266:208-215(1999).
CC -!- TISSUE SPECIFICITY: HEART, NERVOUS SYSTEM, GUT AND KIDNEY.
CC -!- DEVELOPMENTAL STAGE: THE EXPRESSION IS QUITE LOW AT THE STAGE OF
CC 9.5 DPC, BUT BEGAN TO INCREASE AFTER 11.5 DPC.
CC -!- SIMILARITY: BELONGS TO THE NDRG FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB033921; BAA85882.1; -.
DR MGD; MG1:1352498; Ndr2.
DR InterPro; IPR004142; Ndr.
DR Pfam; PF03096; Ndr; 1.
DR SEQUENCE 371 AA; 40789 MW; A5337C04278197B2 CRC64;

```

Query Match	Similarity	50.1%	Score 1038	DB 1	Length 371
Best Local	Similarity	52.1%	Pred. No. 1,76-77		
Matches	208	Conservative	64	Mismatches	89
				Indels	38
				Gaps	77

QY	4	EMQDVLAEXPLV-----	EKGETITGLLEPFDVQEDLETLLGSHVHTLLCTPRKN	55
Db	3	ELQEVQITREKRPILFGQTPRETKAEHLAARILLD-QGQTHSVETGVSTFVTVYTPKPK	61	
QY	56	RPLVLTTHYDDIGMNHNTCYNPLFNEDMDQEIQHFALVCVDPAGOODGAASPAGTMYRSM	115	
Db	62	RPLATFYHDVGLNYSKCFOPLEFRFDMQELLQNFRAVHVDAVGMEEGAPVPLGQYPSL	121	
QY	116	DQLAEMLPGLVLOQFGLKSLIIGMGTGAGAYIILTRFPLNNPMEWEGVLVLTNNPCAGMMDW	175	
Db	122	DQLADMICILIQYLNFSITIIIGVGAGAYILSRVALNHPDVTYEGVLILNIDPNKAGMMDW	181	
QY	176	LAASKTSGTQALPDMVSHLFGKEBMQSNVEVHTYRQHIYNDMNPGLHLFYNAYNSRR	235	
Db	182	AAHKLTGLTSSIPDMILGHLPFSQEBLSQNSBELIQCKRGIIQHPALNLEIETLYMNSYNNRR	241	
QY	236	DLEIERPMPGTHVTVLTQCPALLVVDSDSPAVDAVVEGNSKLDPTTLLTKMADGGGLROI	295	
Db	242	DLNIEER---GGET-TLKCPVMTLVVDQAPHEBDAAVEGNSKLDPTVTSFLKMASSGGQPOL	297	
QY	296	SQPAKLAEAFKYFVQGMGYMPSAMTRLMRSRTHASGSSVTSILDTGRSHRSSEBGRSRSH	355	
Db	298	TQPGKLTAEAFKYFLQGMGYMASSCMTRLSRSRTHALTSAAIDGSRSS-----	346	
QY	356	TSEBGRSHRSSEGAHLDITPRSGAAGNSACKPSMEVSC	394	
Db	347	-----RLTSSQSSSGTLP-----GGPGH-----TWEVSC	371	

```

RESULT 9
SP21 HELAN
ID SP21 HELAN STANDARD; PRT; 352 AA.
AC O23969;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pollen specific protein SP21.
GN SP21.
OS Helianthus annuus (Common sunflower).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; campanulids; Asterales; Asteraceae; Asteroideae;
OC Helianthaceae; Helianthus.
OX NCBI_TaxID=4232;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Flower;
RC Kruter-Caham R., Bronner R., Evrand J.L., Hahne G., Friedt W.,
RA Steimetz A.;
RT "A transmitting tissue- and pollen-expressed protein from sunflower
RL with sequence similarity to the human RTP protein.";
RL Plant Sci. 129:191-202(1997).
CC -1- TISSUE SPECIFICITY: POLLEN.
CC -1- SIMILARITY: BELONGS TO THE NRG FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y09057; CAA70260.1; -.
CC PIR; T13993; T13993.
CC InterPro: IPR004142; Ndr.
CC Pfam; PF03096; Ndr; 1.
CC SEQUENCE 352 AA; 39191 MW; C563B2A225D4497B CRC64;

```

Query Match	17.0%:	Score 351.5:	DB 1:	Length 352:
Best Local Similarity	29.1%:	Pred. No. 2,36-21:		
Matches	92:	Conservative	69:	Mismatches 130: Indels 25: Gaps 8:
QY	16	LVEKGETITGLLOEFVY--QEDDIETLHGSVHTLCTGPKGNRPVILTYHDIGMNHKTCY	73	
Db	1	MADGSHISVTFPFSPFHGSGKEHIIRTCGSGSVTVCDQE--KPDLTTPYDLALNHMSCF	58	
QY	74	NPLF--NEDMGEIIGHFAVCHVDAPGGQDGAASFAGYWRPMOQLMELPGVLDQGGK	132	
Db	59	QGLFVSPESASLLAHNCIYHINPGEHLGASIGIDDPVPSIEDLCQILVLTNYFRLG	118	
QY	133	SIILMGTGAGAYILITRFPALNNPEWVEGLVLINVPKAEWMDMAKIS-----GWTQ	185	
Db	119	SVMGGMAGAYILITLPSIKYSEKRVTLILSPICKAPSWTERFYNKLTSTKTLYYGWCMD	178	
QY	186	ALPBMVYSHLPGEKEMOSNEVHTYR---QHTVNDMPENLLFTINAYNSRDLLEIR	241	
Db	179	LVKSLILHRYFSK--EYCGNPEIPESDYLACRKLIDERSVNVWRVLOALDSRDIREE--	236	
QY	242	PMFSTHTVTLQCPALVVDGSSPAVDAVEGNSKLDPKTKLLTKMAOCGGSPQISOPAKL	301	
Db	237	-----LKSLECKTIIFVGDSSPFHDEALQIAELKGTNCSALVEVHACGSAVVTGEORPHAM	290	
QY	302	AEAFKIVVQGG--YMP	316	
Db	291	LIPLENFLKGFGLYRP	306	

RESULT 10

YDJI CAEEL

ID YDJI CAEEL

STANDARD;

PRT;

325 AA.

002485;

DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical 35.6 kDa protein ZK1073.1 in chromosome X.
 GN ZK1073.1
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 CX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA McKusick A.1;
 RL Submitted (NOV-1995) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE NDRC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; 268135; CAA92227.1; -
 DR PIR; T27688; T27688.
 DR Motmed; ZK1073.1; CE03845.
 DR InterPro; IPR004142; Ndr.
 DR Pfam; PF03096; Ndr; 1.
 DR Hypothetical protein.
 SQ SEQUENCE 325 AA; 35591 MW; F68447B0A4FA74A5 CRC64;

Query Match 16.7%; Score 347; DB 1; Length 325;
 Best Local Similarity 30.4%; Pred. No. 4.8e-21;
 Matches 101; Conservative 55; Mismatches 142; Indels 34; Gaps 11;
 Oy 27 LQEPVQEDIEETLHGSVHTLCTGTPKGRPVILTYHDIGMNHKT-CYNPLFNYEDMOEIT 86
 Db 6 LQMVVVOAGNCGLVAVVQGNL--BERGKTIILTVHDIGTMHKS-FRFPVNHPSMATYK 62
 Oy 87 OHFAVCHVDAPGQDGAASFPAGYVYPSMDQLAEMLPVLAQFGKLSITIGMGTAGAYIL 146
 Db 63 EKAIPLHVCVPGQEDNSADFFGD--FPTLDIGIDDLASVLDKFEVYSALAFEGEGVANI 120
 Oy 147 TRPALNPMVGVILINNPACBGMMDAASKI-----SGTQALPDMVSH 194
 Db 121 CRFAGHPRIIMGIVLHCTSTTAGIIEYCKEKVNMRLSNSIMSDAA--DYLLAH 175
 Oy 195 LFGKEMOSNVENVTVYQHIIVNDNMPGNIHLFINAYNSRDLIERPMPGTHVTLQCP 254
 Db 176 KFGGS-KSRQ---YBELKATINPKLSKIVAFTRKTLDS--STIGKLETVD-- 225
 Oy 255 ALLVVGDSPAVDAYVECNKSLDPTKTLTLKMADCGGLPOISOPAKLAFAFYVQGMGY 314
 Db 226 ALLVVGSKASHLHTVYTHKSNKKKTLVVDNVADVQ--EAPKLAASLILCKGCGV 284
 Oy 315 MFSASMTLRMSRTASGSVSLDGTGRASHT 346
 Db 285 LSGVAIPGMRQRTLS-SSMEADRPBRMSVT 315

RESULT 11
 BIOH_ECOLI
 ID BIOH_ECOLI STANDARD; PRT; 256 AA.
 AC P13001;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE BioH protein.
 GN BIOH OR BIOB OR B3412.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=90016899; PubMed=2678009;
 RA O'Regan M., Gloeckler R., Bernard S., Ledoux C., Ohsawa I.,
 RA Lemoine Y.;
 RT "Nucleotide sequence of the bioH gene of Escherichia coli.";
 RL Nucleic Acids Res. 17:8004-8004(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blatter F.R., Plunkett G. II, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 CC -1- FUNCTION: Not clear. Seems to be implicated in the early steps of
 CC biotin biosynthesis.
 CC -----
 CC -1- PATHWAY: Biotin biosynthesis.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X15587; CAA3612.1; -
 DR EMBL; U18997; AA58210.1; -
 DR EMBL; AB000417; AAC76437.1; -
 DR PIR; Q00081; BVECBH.
 DR PDB; 1M33; 21-JAN-03.
 DR Ecogene; EG10122; bioH.
 DR InterPro; IPR000073; A/B hydrolase.
 DR InterPro; IPR000379; Ser-esterase site.
 DR Pfam; PF00561; abhydrolase_1.
 DR K12 biotin biosynthesis; Complete proteome; 3D-structure.
 SQ SEQUENCE 256 AA; 28505 MW; 93126F241BBCBF3 CRC64;

Query Match 6.7%; Score 138.5; DB 1; Length 256;
 Best Local Similarity 24.2%; Pred. No. 0.0004;
 Matches 69; Conservative 35; Mismatches 80; Indels 101; Gaps 14;
 Oy 42 GSVHTLCTGTPKGRPVILTYHDIGMNHKT--CYNPLFNYEDMOEITOHFAVCHVDAPG- 98
 Db 12 GNVHLVLT-----LHGGLNAEVRICD-----BELSHFTLHVDLPGF 50
 Oy 99 ---QODGAASFPAGYVYPSMDQLAEMLPVLAQFGKLSITIGMGTAGAYILTRPALNPE 155
 Db 51 GRSRGFGMLSS-----LADMEAYVLAQAPDKA-IMVGSISGLVVSQIALTLTRPE 97
 Oy 156 MVEGLVILNVPACBGMMDAASKI-----ISGMTQALPDMVSHLFGKEMOSNVENVTVYR 212
 Db 98 RVQALVTYASSPCCSARDMPGIRKPDVLAGQQQLSD-----DFQRTVE----- 141
 Oy 213 OHIVNDNMPGNIHLFINAYNSRDLIERPMPGTHVTLQCPAL--LVVGDSSPAVDAYV 270
 Db 142 -----RFLAQ--TMGETETANQDABALKVTLALPMPREVTVL- 176
 Oy 271 EONSKLDPKTLTLKMAADCGLPOISOPAKLAFAFYVQGMGY 315
 Db 177 --NGLEIKLVTDLRQF-----LQNVSMF-----FLRYGYL 206

RESULT 12
 TENA_CHICK
 ID TENA_CHICK STANDARD; PRT; 1808 AA.
 AC P10039; O73584; O73585; P11132;
 DT 01-MAR-1989 (Rel. 10, Created)

DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Tenascin precursor (TN) (Hexabractin) (Cytotactin) (Neuronectin)
 DE (GHEM) (U1) (Mitochondrial antigen) (Glioma-associated extracellular
 DE matrix antigen) (GP 150-225).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NC NCBI_TaxID=9031;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA MEDLINE=90030407; PubMed=2478295;
 RA Spring J., Beck K., Chiquet-Ehrismann R.;
 RT "Two contrary functions of tenascin: dissection of the active sites
 RT by recombinant tenascin fragments.";
 RL Cell 59:325-334(1989).
 RN [2]
 RN SEQUENCE OF 27-722 FROM N.A., AND SEQUENCE OF 79-96.
 RC TISSUE=Fibroblast;
 RA MEDLINE=89030589; PubMed=2460335;
 RA Pearson C.A., Pearson D., Shibahara S., Hofsteenge J.,
 RA Chiquet-Ehrismann R.;
 RT "Tenascin: cDNA cloning and induction by TGF-beta.";
 RL EMBO J. 7:2977-2982(1988).
 RN [3]
 RN SEQUENCE OF 464-1018 AND 1412-1661 FROM N.A., AND SEQUENCE OF 852-868.
 RC TISSUE=Embryo;
 RA MEDLINE=88176910; PubMed=2451243;
 RA Jones F.S., Burgon M.P., Hoffman S., Crossin K.L., Cunningham B.A.,
 RA Edelman G.M.;
 RT "A cDNA clone for cytotactin contains sequences similar to epidermal
 RT growth factor-like repeats and segments of fibronectin and
 RT fibronogen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:2186-2190(1988).
 CC -1- FUNCTION: SAM (SUBSTRATE-ADHESION MOLECULE) THAT APPEARS TO
 CC INHIBIT CELL MIGRATION. MAY PLAY A ROLE IN SUPPORTING THE GROWTH
 CC OF EPITHELIAL TUMORS. IS A LIGAND FOR INTEGRINS ALPHA-8/BETA-1,
 CC ALPHA-9/BETA-1, ALPHA-V/BETA-3 AND ALPHA-V/BETA-6.
 CC -1- SUBUNIT: HEXAMERIC. A HOMOTRIMER MAY BE FORMED IN THE TRIPLE
 CC COILED-COIL REGION AND MAY BE STABILIZED BY DISULFIDE RINGS AT
 CC BOTH ENDS. TWO OF SUCH HALF-HEXAMER ACTIONS MAY BE DISULFIDE LINKED
 CC WITHIN THE CENTRAL GLOBULE.
 CC -1- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
 CC -1- ALTERNATIVE PRODUCTS: Secreted; extracellular matrix.
 CC Event=Alternative splicing; Named isoforms=3;
 CC Comment=Isoforms are produced in a tissue- and time-specific
 CC manner during development;
 CC Name=1; Synonyms=230 kDa;
 CC IsoId=P10039-1; Sequence=Displayed;
 CC Name=2; Synonyms=200 kDa;
 CC IsoId=P10039-2; Sequence=VSP_001410;
 CC Name=3; Synonyms=190 kDa;
 CC IsoId=P10039-3; Sequence=VSP_001411;
 CC -1- INDUCTION: By TGF-beta.
 CC -1- SIMILARITY: Contains 14 EGF-like domains.
 CC -1- SIMILARITY: Contains 1 fibronectin type III domains.
 CC -1- SIMILARITY: Contains 1 fibronectin C-terminal domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL; M23121; AAA49086.1; -;
 DR EMBL; X08031; CAB40811.1; -;
 DR EMBL; X08030; CAB40824.1; ALT_TERM.
 DR EMBL; J03641; AAA48748.1; ALT_SEQ.
 DR EMBL; M20816; AAA48749.1; ALT_SEQ.

DR PDB; 1OR4; 29-DEC-99.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR002181; Fibronogen_C.
 DR InterPro; IPR003961; FN III.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR002049; Lamitin_EGF.
 DR Pfam; PF00008; EGF; 13.
 DR Pfam; PF00147; fibronogen_C; 1.
 DR Pfam; PF00041; fn3; 11.
 DR PRINTS; PR00011; EGF-LAMININ.
 DR SMART; SM00181; EGF; 10.
 DR SMART; SM00186; FBG; 1.
 DR SMART; SM00060; FN3; 9.
 DR PROSITE; PS00022; EGF_1; 14.
 DR PROSITE; PS01166; EGF_2; 14.
 DR Glycoprotein; Cell adhesion; Repeat; EGF-like domain; Coiled coil;
 DR Extracellular matrix; Signal; Alternative splicing; 3D-structure.
 KW SIGNAL
 FT 1
 FT PROPEP 23 33
 FT CHAIN 34 1808
 FT DOMAIN 116 142
 FT DOMAIN 176 188
 FT DOMAIN 188 219
 FT DOMAIN 219 250
 FT DOMAIN 250 281
 FT DOMAIN 281 312
 FT DOMAIN 312 343
 FT DOMAIN 343 374
 FT DOMAIN 374 405
 FT DOMAIN 405 436
 FT DOMAIN 436 467
 FT DOMAIN 467 498
 FT DOMAIN 498 529
 FT DOMAIN 529 560
 FT DOMAIN 560 591
 FT DOMAIN 591 680
 FT DOMAIN 681 771
 FT DOMAIN 772 862
 FT DOMAIN 863 954
 FT DOMAIN 955 1042
 FT DOMAIN 1043 1133
 FT DOMAIN 1134 1224
 FT DOMAIN 1225 1315
 FT DOMAIN 1316 1404
 FT DOMAIN 1405 1492
 FT DOMAIN 1493 1580
 FT DOMAIN 1589 1808
 FT DOMAIN 64 64
 FT DISULFID 192 202
 FT DISULFID 196 207
 FT DISULFID 209 218
 FT DISULFID 223 233
 FT DISULFID 227 238
 FT DISULFID 240 249
 FT DISULFID 254 264
 FT DISULFID 258 269
 FT DISULFID 271 280
 FT DISULFID 285 295
 FT DISULFID 289 300
 FT DISULFID 302 311
 FT DISULFID 316 326
 FT DISULFID 320 331
 FT DISULFID 333 342
 FT DISULFID 347 357
 FT DISULFID 351 362
 FT DISULFID 364 373
 FT DISULFID 378 388
 FT DISULFID 382 393
 FT DISULFID 395 404
 FT DISULFID 409 419
 FT DISULFID 413 424
 FT DISULFID 426 435
 FT DISULFID 440 450
 FT DISULFID 440 450
 TENASCIN.
 COILED COIL (POTENTIAL).
 EGF-LIKE 1 (INCOMPLETE).
 EGF-LIKE 2.
 EGF-LIKE 3.
 EGF-LIKE 4.
 EGF-LIKE 5.
 EGF-LIKE 6.
 EGF-LIKE 7.
 EGF-LIKE 8.
 EGF-LIKE 9.
 EGF-LIKE 10.
 EGF-LIKE 11.
 EGF-LIKE 12.
 EGF-LIKE 13.
 EGF-LIKE 14.
 FIBRONECTIN TYPE-III 1.
 FIBRONECTIN TYPE-III 2.
 FIBRONECTIN TYPE-III 3.
 FIBRONECTIN TYPE-III 4.
 FIBRONECTIN TYPE-III 5.
 FIBRONECTIN TYPE-III 6.
 FIBRONECTIN TYPE-III 7.
 FIBRONECTIN TYPE-III 8.
 FIBRONECTIN TYPE-III 9.
 FIBRONECTIN TYPE-III 10.
 FIBRONECTIN TYPE-III 11.
 FIBRONECTIN TYPE-III 12.
 FIBRONECTIN TYPE-III 13.
 FIBRONECTIN TYPE-III 14.
 FIBRONECTIN TYPE-III 15.
 FIBRONECTIN TYPE-III 16.
 FIBRONECTIN TYPE-III 17.
 FIBRONECTIN TYPE-III 18.
 FIBRONECTIN TYPE-III 19.
 FIBRONECTIN TYPE-III 20.
 FIBRONECTIN TYPE-III 21.
 FIBRONECTIN TYPE-III 22.
 FIBRONECTIN TYPE-III 23.
 FIBRONECTIN TYPE-III 24.
 FIBRONECTIN TYPE-III 25.
 FIBRONECTIN TYPE-III 26.
 FIBRONECTIN TYPE-III 27.
 FIBRONECTIN TYPE-III 28.
 FIBRONECTIN TYPE-III 29.
 FIBRONECTIN TYPE-III 30.
 FIBRONECTIN TYPE-III 31.
 FIBRONECTIN TYPE-III 32.
 FIBRONECTIN TYPE-III 33.
 FIBRONECTIN TYPE-III 34.
 FIBRONECTIN TYPE-III 35.
 FIBRONECTIN TYPE-III 36.
 FIBRONECTIN TYPE-III 37.
 FIBRONECTIN TYPE-III 38.
 FIBRONECTIN TYPE-III 39.
 FIBRONECTIN TYPE-III 40.
 FIBRONECTIN TYPE-III 41.
 FIBRONECTIN TYPE-III 42.
 FIBRONECTIN TYPE-III 43.
 FIBRONECTIN TYPE-III 44.
 FIBRONECTIN TYPE-III 45.
 FIBRONECTIN TYPE-III 46.
 FIBRONECTIN TYPE-III 47.
 FIBRONECTIN TYPE-III 48.
 FIBRONECTIN TYPE-III 49.
 FIBRONECTIN TYPE-III 50.
 FIBRONECTIN TYPE-III 51.
 FIBRONECTIN TYPE-III 52.
 FIBRONECTIN TYPE-III 53.
 FIBRONECTIN TYPE-III 54.
 FIBRONECTIN TYPE-III 55.
 FIBRONECTIN TYPE-III 56.
 FIBRONECTIN TYPE-III 57.
 FIBRONECTIN TYPE-III 58.
 FIBRONECTIN TYPE-III 59.
 FIBRONECTIN TYPE-III 60.
 FIBRONECTIN TYPE-III 61.
 FIBRONECTIN TYPE-III 62.
 FIBRONECTIN TYPE-III 63.
 FIBRONECTIN TYPE-III 64.
 FIBRONECTIN TYPE-III 65.
 FIBRONECTIN TYPE-III 66.
 FIBRONECTIN TYPE-III 67.
 FIBRONECTIN TYPE-III 68.
 FIBRONECTIN TYPE-III 69.
 FIBRONECTIN TYPE-III 70.
 FIBRONECTIN TYPE-III 71.
 FIBRONECTIN TYPE-III 72.
 FIBRONECTIN TYPE-III 73.
 FIBRONECTIN TYPE-III 74.
 FIBRONECTIN TYPE-III 75.
 FIBRONECTIN TYPE-III 76.
 FIBRONECTIN TYPE-III 77.
 FIBRONECTIN TYPE-III 78.
 FIBRONECTIN TYPE-III 79.
 FIBRONECTIN TYPE-III 80.
 FIBRONECTIN TYPE-III 81.
 FIBRONECTIN TYPE-III 82.
 FIBRONECTIN TYPE-III 83.
 FIBRONECTIN TYPE-III 84.
 FIBRONECTIN TYPE-III 85.
 FIBRONECTIN TYPE-III 86.
 FIBRONECTIN TYPE-III 87.
 FIBRONECTIN TYPE-III 88.
 FIBRONECTIN TYPE-III 89.
 FIBRONECTIN TYPE-III 90.
 FIBRONECTIN TYPE-III 91.
 FIBRONECTIN TYPE-III 92.
 FIBRONECTIN TYPE-III 93.
 FIBRONECTIN TYPE-III 94.
 FIBRONECTIN TYPE-III 95.
 FIBRONECTIN TYPE-III 96.
 FIBRONECTIN TYPE-III 97.
 FIBRONECTIN TYPE-III 98.
 FIBRONECTIN TYPE-III 99.
 FIBRONECTIN TYPE-III 100.
 FIBRONECTIN TYPE-III 101.
 FIBRONECTIN TYPE-III 102.
 FIBRONECTIN TYPE-III 103.
 FIBRONECTIN TYPE-III 104.
 FIBRONECTIN TYPE-III 105.
 FIBRONECTIN TYPE-III 106.
 FIBRONECTIN TYPE-III 107.
 FIBRONECTIN TYPE-III 108.
 FIBRONECTIN TYPE-III 109.
 FIBRONECTIN TYPE-III 110.
 FIBRONECTIN TYPE-III 111.
 FIBRONECTIN TYPE-III 112.
 FIBRONECTIN TYPE-III 113.
 FIBRONECTIN TYPE-III 114.
 FIBRONECTIN TYPE-III 115.
 FIBRONECTIN TYPE-III 116.
 FIBRONECTIN TYPE-III 117.
 FIBRONECTIN TYPE-III 118.
 FIBRONECTIN TYPE-III 119.
 FIBRONECTIN TYPE-III 120.
 FIBRONECTIN TYPE-III 121.
 FIBRONECTIN TYPE-III 122.
 FIBRONECTIN TYPE-III 123.
 FIBRONECTIN TYPE-III 124.
 FIBRONECTIN TYPE-III 125.
 FIBRONECTIN TYPE-III 126.
 FIBRONECTIN TYPE-III 127.
 FIBRONECTIN TYPE-III 128.
 FIBRONECTIN TYPE-III 129.
 FIBRONECTIN TYPE-III 130.
 FIBRONECTIN TYPE-III 131.
 FIBRONECTIN TYPE-III 132.
 FIBRONECTIN TYPE-III 133.
 FIBRONECTIN TYPE-III 134.
 FIBRONECTIN TYPE-III 135.
 FIBRONECTIN TYPE-III 136.
 FIBRONECTIN TYPE-III 137.
 FIBRONECTIN TYPE-III 138.
 FIBRONECTIN TYPE-III 139.
 FIBRONECTIN TYPE-III 140.
 FIBRONECTIN TYPE-III 141.
 FIBRONECTIN TYPE-III 142.
 FIBRONECTIN TYPE-III 143.
 FIBRONECTIN TYPE-III 144.
 FIBRONECTIN TYPE-III 145.
 FIBRONECTIN TYPE-III 146.
 FIBRONECTIN TYPE-III 147.
 FIBRONECTIN TYPE-III 148.
 FIBRONECTIN TYPE-III 149.
 FIBRONECTIN TYPE-III 150.
 FIBRONECTIN TYPE-III 151.
 FIBRONECTIN TYPE-III 152.
 FIBRONECTIN TYPE-III 153.
 FIBRONECTIN TYPE-III 154.
 FIBRONECTIN TYPE-III 155.
 FIBRONECTIN TYPE-III 156.
 FIBRONECTIN TYPE-III 157.
 FIBRONECTIN TYPE-III 158.
 FIBRONECTIN TYPE-III 159.
 FIBRONECTIN TYPE-III 160.
 FIBRONECTIN TYPE-III 161.
 FIBRONECTIN TYPE-III 162.
 FIBRONECTIN TYPE-III 163.
 FIBRONECTIN TYPE-III 164.
 FIBRONECTIN TYPE-III 165.
 FIBRONECTIN TYPE-III 166.
 FIBRONECTIN TYPE-III 167.
 FIBRONECTIN TYPE-III 168.
 FIBRONECTIN TYPE-III 169.
 FIBRONECTIN TYPE-III 170.
 FIBRONECTIN TYPE-III 171.
 FIBRONECTIN TYPE-III 172.
 FIBRONECTIN TYPE-III 173.
 FIBRONECTIN TYPE-III 174.
 FIBRONECTIN TYPE-III 175.
 FIBRONECTIN TYPE-III 176.
 FIBRONECTIN TYPE-III 177.
 FIBRONECTIN TYPE-III 178.
 FIBRONECTIN TYPE-III 179.
 FIBRONECTIN TYPE-III 180.
 FIBRONECTIN TYPE-III 181.
 FIBRONECTIN TYPE-III 182.
 FIBRONECTIN TYPE-III 183.
 FIBRONECTIN TYPE-III 184.
 FIBRONECTIN TYPE-III 185.
 FIBRONECTIN TYPE-III 186.
 FIBRONECTIN TYPE-III 187.
 FIBRONECTIN TYPE-III 188.
 FIBRONECTIN TYPE-III 189.
 FIBRONECTIN TYPE-III 190.
 FIBRONECTIN TYPE-III 191.
 FIBRONECTIN TYPE-III 192.
 FIBRONECTIN TYPE-III 193.
 FIBRONECTIN TYPE-III 194.
 FIBRONECTIN TYPE-III 195.
 FIBRONECTIN TYPE-III 196.
 FIBRONECTIN TYPE-III 197.
 FIBRONECTIN TYPE-III 198.
 FIBRONECTIN TYPE-III 199.
 FIBRONECTIN TYPE-III 200.
 FIBRONECTIN TYPE-III 201.
 FIBRONECTIN TYPE-III 202.
 FIBRONECTIN TYPE-III 203.
 FIBRONECTIN TYPE-III 204.
 FIBRONECTIN TYPE-III 205.
 FIBRONECTIN TYPE-III 206.
 FIBRONECTIN TYPE-III 207.
 FIBRONECTIN TYPE-III 208.
 FIBRONECTIN TYPE-III 209.
 FIBRONECTIN TYPE-III 210.
 FIBRONECTIN TYPE-III 211.
 FIBRONECTIN TYPE-III 212.
 FIBRONECTIN TYPE-III 213.
 FIBRONECTIN TYPE-III 214.
 FIBRONECTIN TYPE-III 215.
 FIBRONECTIN TYPE-III 216.
 FIBRONECTIN TYPE-III 217.
 FIBRONECTIN TYPE-III 218.
 FIBRONECTIN TYPE-III 219.
 FIBRONECTIN TYPE-III 220.
 FIBRONECTIN TYPE-III 221.
 FIBRONECTIN TYPE-III 222.
 FIBRONECTIN TYPE-III 223.
 FIBRONECTIN TYPE-III 224.
 FIBRONECTIN TYPE-III 225.
 FIBRONECTIN TYPE-III 226.
 FIBRONECTIN TYPE-III 227.
 FIBRONECTIN TYPE-III 228.
 FIBRONECTIN TYPE-III 229.
 FIBRONECTIN TYPE-III 230.
 FIBRONECTIN TYPE-III 231.
 FIBRONECTIN TYPE-III 232.
 FIBRONECTIN TYPE-III 233.
 FIBRONECTIN TYPE-III 234.
 FIBRONECTIN TYPE-III 235.
 FIBRONECTIN TYPE-III 236.
 FIBRONECTIN TYPE-III 237.
 FIBRONECTIN TYPE-III 238.
 FIBRONECTIN TYPE-III 239.
 FIBRONECTIN TYPE-III 240.
 FIBRONECTIN TYPE-III 241.
 FIBRONECTIN TYPE-III 242.
 FIBRONECTIN TYPE-III 243.
 FIBRONECTIN TYPE-III 244.
 FIBRONECTIN TYPE-III 245.
 FIBRONECTIN TYPE-III 246.
 FIBRONECTIN TYPE-III 247.
 FIBRONECTIN TYPE-III 248.
 FIBRONECTIN TYPE-III 249.
 FIBRONECTIN TYPE-III 250.
 FIBRONECTIN TYPE-III 251.
 FIBRONECTIN TYPE-III 252.
 FIBRONECTIN TYPE-III 253.
 FIBRONECTIN TYPE-III 254.
 FIBRONECTIN TYPE-III 255.
 FIBRONECTIN TYPE-III 256.
 FIBRONECTIN TYPE-III 257.
 FIBRONECTIN TYPE-III 258.
 FIBRONECTIN TYPE-III 259.
 FIBRONECTIN TYPE-III 260.
 FIBRONECTIN TYPE-III 261.
 FIBRONECTIN TYPE-III 262.
 FIBRONECTIN TYPE-III 263.
 FIBRONECTIN TYPE-III 264.
 FIBRONECTIN TYPE-III 265.
 FIBRONECTIN TYPE-III 266.
 FIBRONECTIN TYPE-III 267.
 FIBRONECTIN TYPE-III 268.
 FIBRONECTIN TYPE-III 269.
 FIBRONECTIN TYPE-III 270.
 FIBRONECTIN TYPE-III 271.
 FIBRONECTIN TYPE-III 272.
 FIBRONECTIN TYPE-III 273.
 FIBRONECTIN TYPE-III 274.
 FIBRONECTIN TYPE-III 275.
 FIBRONECTIN TYPE-III 276.
 FIBRONECTIN TYPE-III 277.
 FIBRONECTIN TYPE-III 278.
 FIBRONECTIN TYPE-III 279.
 FIBRONECTIN TYPE-III 280.
 FIBRONECTIN TYPE-III 281.
 FIBRONECTIN TYPE-III 282.
 FIBRONECTIN TYPE-III 283.
 FIBRONECTIN TYPE-III 284.
 FIBRONECTIN TYPE-III 285.
 FIBRONECTIN TYPE-III 286.
 FIBRONECTIN TYPE-III 287.
 FIBRONECTIN TYPE-III 288.
 FIBRONECTIN TYPE-III 289.
 FIBRONECTIN TYPE-III 290.
 FIBRONECTIN TYPE-III 291.
 FIBRONECTIN TYPE-III 292.
 FIBRONECTIN TYPE-III 293.
 FIBRONECTIN TYPE-III 294.
 FIBRONECTIN TYPE-III 295.
 FIBRONECTIN TYPE-III 296.
 FIBRONECTIN TYPE-III 297.
 FIBRONECTIN TYPE-III 298.
 FIBRONECTIN TYPE-III 299.
 FIBRONECTIN TYPE-III 300.
 FIBRONECTIN TYPE-III 301.
 FIBRONECTIN TYPE-III 302.
 FIBRONECTIN TYPE-III 303.
 FIBRONECTIN TYPE-III 304.
 FIBRONECTIN TYPE-III 305.
 FIBRONECTIN TYPE-III 306.
 FIBRONECTIN TYPE-III 307.
 FIBRONECTIN TYPE-III 308.
 FIBRONECTIN TYPE-III 309.
 FIBRONECTIN TYPE-III 310.
 FIBRONECTIN TYPE-III 311.
 FIBRONECTIN TYPE-III 312.
 FIBRONECTIN TYPE-III 313.
 FIBRONECTIN TYPE-III 314.
 FIBRONECTIN TYPE-III 315.
 FIBRONECTIN TYPE-III 316.
 FIBRONECTIN TYPE-III 317.
 FIBRONECTIN TYPE-III 318.
 FIBRONECTIN TYPE-III 319.
 FIBRONECTIN TYPE-III 320.
 FIBRONECTIN TYPE-III 321.
 FIBRONECTIN TYPE-III 322.
 FIBRONECTIN TYPE-III 323.
 FIBRONECTIN TYPE-III 324.
 FIBRONECTIN TYPE-III 325.
 FIBRONECTIN TYPE-III 326.
 FIBRONECTIN TYPE-III 327.
 FIBRONECTIN TYPE-III 328.
 FIBRONECTIN TYPE-III 329.
 FIBRONECTIN TYPE-III 330.
 FIBRONECTIN TYPE-III 331.
 FIBRONECTIN TYPE-III 332.
 FIBRONECTIN TYPE-III 333.
 FIBRONECTIN TYPE-III 334.
 FIBRONECTIN TYPE-III 335.
 FIBRONECTIN TYPE-III 336.
 FIBRONECTIN TYPE-III 337.
 FIBRONECTIN TYPE-III 338.
 FIBRONECTIN TYPE-III 339.
 FIBRONECTIN TYPE-III 340.
 FIBRONECTIN TYPE-III 341.
 FIBRONECTIN TYPE-III 342.
 FIBRONECTIN TYPE-III 343.
 FIBRONECTIN TYPE-III 344.
 FIBRONECTIN TYPE-III 345.
 FIBRONECTIN TYPE-III 346.
 FIBRONECTIN TYPE-III 347.
 FIBRONECTIN TYPE-III 348.
 FIBRONECTIN TYPE-III 349.
 FIBRONECTIN TYPE-III 350.
 FIBRONECTIN TYPE-III 351.
 FIBRONECTIN TYPE-III 352.
 FIBRONECTIN TYPE-III 353.
 FIBRONECTIN TYPE-III 354.
 FIBRONECTIN TYPE-III 355.
 FIBRONECTIN TYPE-III 356.
 FIBRONECTIN TYPE-III 357.
 FIBRONECTIN TYPE-III 358.
 FIBRONECTIN TYPE-III 359.
 FIBRONECTIN TYPE-III 360.
 FIBRONECTIN TYPE-III 361.
 FIBRONECTIN TYPE-III 362.
 FIBRONECTIN TYPE-III 363.
 FIBRONECTIN TYPE-III 364.
 FIBRONECTIN TYPE-III 365.
 FIBRONECTIN TYPE-III 366.
 FIBRONECTIN TYPE-III 367.
 FIBRONECTIN TYPE-III 368.
 FIBRONECTIN TYPE-III 369.
 FIBRONECTIN TYPE-III 370.
 FIBRONECTIN TYPE-III 371.
 FIBRONECTIN TYPE-III 372.
 FIBRONECTIN TYPE-III 373.
 FIBRONECTIN TYPE-III 374.
 FIBRONECTIN TYPE-III 375.
 FIBRONECTIN TYPE-III 376.
 FIBRONECTIN TYPE-III 377.
 FIBRONECTIN TYPE-III 378.
 FIBRONECTIN TYPE-III 379.
 FIBRONECTIN TYPE-III 380.
 FIBRONECTIN TYPE-III 381.
 FIBRONECTIN TYPE-III 382.
 FIBRONECTIN TYPE-III 383.
 FIBRONECTIN TYPE-III 384.
 FIBRONECTIN TYPE-III 385.
 FIBRONECTIN TYPE-III 386.
 FIBRONECTIN TYPE-III 387.
 FIBRONECTIN TYPE-III 388.
 FIBRONECTIN TYPE-III 389.
 FIBRONECTIN TYPE-III 390.
 FIBRONECTIN TYPE-III 391.
 FIBRONECTIN TYPE-III 392.
 FIBRONECTIN TYPE-III 393.
 FIBRONECTIN TYPE-III 394.
 FIBRONECTIN TYPE-III 395.
 FIBRONECTIN TYPE-III 396.
 FIBRONECTIN TYPE-III 397.
 FIBRONECTIN TYPE-III 398.
 FIBRONECTIN TYPE-III 399.
 FIBRONECTIN TYPE-III 400.
 FIBRONECTIN TYPE-III 401.
 FIBRONECTIN TYPE-III 402.
 FIBRONECTIN TYPE-III 403.
 FIBRONECTIN TYPE-III 404.
 FIBRONECTIN TYPE-III 405.
 FIBRONECTIN TYPE-III 406.
 FIBRONECTIN TYPE-III 407.
 FIBRONECTIN TYPE-III 408.
 FIBRONECTIN TYPE-III 409.
 FIBRONECTIN TYPE-III 410.
 FIBRONECTIN TYPE-III 411.
 FIBRONECTIN TYPE-III 412.
 FIBRONECTIN TYPE-III 413.
 FIBRONECTIN TYPE-III 414.
 FIBRONECTIN TYPE-III 415.
 FIBRONECTIN TYPE-III 416.
 FIBRONECTIN TYPE-III 417.
 FIBRONECTIN TYPE-III 418.
 FIBRONECTIN TYPE-III 419.
 FIBRONECTIN TYPE-III 420.
 FIBRONECTIN TYPE-III 421.
 FIBRONECTIN TYPE-III 422.
 FIBRONECTIN TYPE-III 423.
 FIBRONECTIN TYPE-III 424.
 FIBRONECTIN TYPE-III 425.
 FIBRONECTIN TYPE-III 426.
 FIBRONECTIN TYPE-III 427.
 FIBRONECTIN TYPE-III 428.
 FIBRONECTIN TYPE-III 429.
 FIBRONECTIN TYPE-III 430.
 FIBRONECTIN TYPE-III 431.
 FIBRONECTIN TYPE-III 432.
 FIBRONECTIN TYPE-III 433.
 FIBRONECTIN TYPE-III 434.
 FIBRONECTIN TYPE-III 435.
 FIBRONECTIN TYPE-III 436.
 FIBRONECTIN TYPE-III 437.
 FIBRONECTIN TYPE-III 438.
 FIBRONECTIN TYPE-III 439.
 FIBRONECTIN TYPE-III 440.
 FIBRONECTIN TYPE-III 441.
 FIBRONECTIN TYPE-III 442.
 FIBRONECTIN TYPE-III 443.
 FIBRONECTIN TYPE-III 444.
 FIBRONECTIN TYPE-III 445.
 FIBRONECTIN TYPE-III 446.
 FIBRONECTIN TYPE-III 447.
 FIBRONECTIN TYPE-III 448.
 FIBRONECTIN TYPE-III 449.
 FIBRONECTIN TYPE-III 450.
 FIBRONECTIN TYPE-III 451.
 FIBRONECTIN TYPE-III 452.
 FIBRONECTIN TYPE-III 453.
 FIBRONECTIN TYPE-III 454.
 FIBRONECTIN TYPE-III 455.
 FIBRONECTIN TYPE-III 456.
 FIBRONECTIN TYPE-III 457.
 FIBRONECTIN TYPE-III 458.
 FIBRONECTIN TYPE-III 459.
 FIBRONECTIN TYPE-III 460.
 FIBRONECTIN TYPE-III 461.
 FIBRONECTIN TYPE-III 462.
 FIBRONECTIN TYPE-III 463.
 FIBRONECTIN TYPE-III 464.
 FIBRONECTIN TYPE-III 465.
 FIBRONECTIN TYPE-III 466.
 FIBRONECTIN TYPE-III 467.
 FIBRONECTIN TYPE-III 468.
 FIBRONECTIN TYPE-III 469.
 FIBRONECTIN TYPE-III 470.
 FIBRONECTIN TYPE-III 471.
 FIBRONECTIN TYPE-III 472.
 FIBRONECTIN TYPE-III 473.
 FIBRONECTIN TYPE-III 474.
 FIBRONECTIN TYPE-III 475.
 FIBRONECTIN TYPE-III 476.
 FIBRONECTIN TYPE-III 477.
 FIBRONECTIN TYPE-III 478.
 FIBRONECTIN TYPE-III 479.
 FIBRONECTIN TYPE-III 480.
 FIBRONECTIN TYPE-III 481.
 FIBRONECTIN TYPE-III 482.
 FIBRONECTIN TYPE-III 483.
 FIBRONECTIN TYPE-III 484.
 FIBRONECTIN TYPE-III 485.
 FIBRONECTIN TYPE-III 486.
 FIBRONECTIN TYPE-III 487.
 FIBRONECTIN TYPE-III 488.
 FIBRONECTIN TYPE-III 489.
 FIBRONECTIN TYPE-III 490.
 FIBRONECTIN TYPE-III 491.
 FIBRONECTIN TYPE-III 492.
 FIBRONECTIN TYPE-III 493.
 FIBRONECTIN TYPE-III 494.
 FIBRONECTIN TYPE-III 495.
 FIBRONECTIN TYPE-III 496.
 FIBRONECTIN TYPE-III 497.
 FIBRONECTIN TYPE-III 498.
 FIBRONECTIN TYPE-III 499.
 FIBRONECTIN TYPE-III 500.
 FIBRONECTIN TYPE-III 501.
 FIBRONECTIN TYPE-III 502.
 FIBRONECTIN TYPE-III 503.
 FIBRONECTIN TYPE-III 504.
 FIBRONECTIN TYPE-III 505.
 FIBRONECTIN TYPE-III 506.
 FIBRONECTIN TYPE-III 507.
 FIBRONECTIN TYPE-III 508.
 FIBRONECTIN TYPE-III 509.
 FIBRONECTIN TYPE-III 510.
 FIBRONECTIN TYPE-III 511.
 FIBRONECTIN TYPE-III 512.
 FIBRONECTIN TYPE-III 513.
 FIBRONECTIN TYPE-III 514.
 FIBRONECTIN TYPE-III 515.
 FIBRONECTIN TYPE-III 516.
 FIBRONECTIN TYPE

```

FT DISULFID 444 455 BY SIMILARITY.
FT DISULFID 457 466 BY SIMILARITY.
FT DISULFID 471 481 BY SIMILARITY.
FT DISULFID 475 486 BY SIMILARITY.
FT DISULFID 488 497 BY SIMILARITY.
FT DISULFID 502 512 BY SIMILARITY.
FT DISULFID 506 517 BY SIMILARITY.
FT DISULFID 519 528 BY SIMILARITY.
FT DISULFID 533 543 BY SIMILARITY.
FT DISULFID 537 548 BY SIMILARITY.
FT DISULFID 550 559 BY SIMILARITY.
FT DISULFID 564 574 BY SIMILARITY.
FT DISULFID 568 579 BY SIMILARITY.
FT DISULFID 581 590 BY SIMILARITY.
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 603 603 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 643 643 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 751 751 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 759 759 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1050 1050 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1090 1090 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1101 1101 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1112 1112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1153 1153 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1183 1183 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1416 1416 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1736 1736 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1769 1769 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 1043 1224 Missing (in isoform 2).
FT VARSPLIC 1043 1315 Missing (in isoform 3).
FT VARSPLIC 1043 1315 /FtId=VSP_001411.
FT CONFLICT 182 182 W -> R (IN REF. 2).
FT CONFLICT 553 571 SCPNDCNNV -> PAMPATTW (IN REF. 3).
FT CONFLICT 598 598 E -> G (IN REF. 3).
FT CONFLICT 838 838 T -> TEY (IN REF. 3).
FT CONFLICT 886 886 N -> F (IN REF. 3).
SQ SEQUENCE 1808 AA; 198858 MW; B924A06CF9FDD5E CRC64;

Query Match
Best Local Similarity 5.2%; Score 107.5; DB 1; Length 1808;
Matches 75; Conservative 61; Mismatches 153; Indels 95; Gaps 18;

13 VKPLVEKGETTGLQ--EFDVQEDITELG-----SVHVLCTGPKGNRPVI 59
1274 IVPGERHTDITGLKEGEYEIE-----LYGVSSGRSROPINSVATTVGSFKG----- 1322
60 LTYHIDGMHKTCTNPLFRYEDMOETQHFANVCHYDARQODGAASFPGNYRPSMDOLA 119
1333 ISFSITENSAT-----VSWTPRSRVDYSRYVPIGTGTNNVTVDGS--KTRTKLV 1374
120 EMLPGLVLOFGKSIIGMGTGAGAYILTRPALNPEWEGVLIVNPPCAGEMDMAASK 179
1375 KLVPEVDVNNVNIISVKGFESEPISIGILKTALDSP---SGLVVNMITD--SEALATQQA- 1429
180 ISGWTQALPDVNVSHLFGKEKQSNVEV-----HTYRQIIVDMNPG 222
1430 ----IAAVDNYIVS--YSESEDEPVTQWVSGNTVEYDINGLARPAETEYTLRVAIVDAQKS 1483
223 NL--HLFINAVNSRDLLEIRPMGTHVTLQCPALLVVGSSPAVDAVECNS----- 274
1484 ETLSTQFTGLDAPDLATEVQSESTAVITWRP-----RAPVTDLILTYESIDGRVK 1536
275 ---KLDP--TKTTLKMAADCGGLPOISQAKLAFAF-----YFVGQMGY-MPSAS 319
1537 EYILDPETTSYTLTELS-----PSTQYVYKQALSRMSRSMKIQTIVFTTGLIYPPKDC 1591
320 MTRLMSRTASGSSVTSLDGTRSR 343
1592 SQALLNGEVTSGLYTIYINGDRTO 1615

```

```

RESULT 13
ID ILVC_SYNEI STANDARD; PRT; 331 AA.
AC Q8DGR0;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE ketol-acid reductoisomerase (EC 1.1.1.86) (Acetohydroxy-acid
DE isomeroeductase) (Alpha-Keto-beta-hydroxyacil reductoisomerase).
GN ILVC OR TL12254.
OS Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxId=32046;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=BP-1;
RX MEDLINE=22225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1."
RL DNA Res. 9:133-130(2002)
CC -I- CATALYTIC ACTIVITY: (R)-2,3-dihydroxy-3-methylbutanoate + NADP(+)
CC = (S)-2-hydroxy-2-methyl-3-oxobutanoate + NADPH.
CC -I- PATHWAY: Valine and isoleucine biosynthesis; second step.
CC -I- SIMILARITY: Belongs to the ketol-acid reductoisomerase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AP005376; BAC09806.1; ALT_INIT.
CC HAMAP: MF_00435; -; 1.
CC InterPro: IPR000506; ACh_isomrdctse.
CC Pfam: PF01450; ilvc; 1.
CC TIGRFAMs: TIGR00465; ilvc; 1.
CC Oxidoreductase; Branched-chain amino acid biosynthesis; NADP;
CC Complete proteome.
CC ACT SITE 108 108 POTENTIAL.
CC SEQUENCE 331 AA; 36403 MW; E1E162D0B57D375 CRC64;

Query Match
Best Local Similarity 4.8%; Score 99.5; DB 1; Length 331;
Matches 63; Conservative 36; Mismatches 96; Indels 85; Gaps 14;

126 LQOFGKSIIGMGTGAGAYILTRPALNPEWEGVLIVNPPCAGEMDMAASKISGTQ 185
36 LRDSGLANNVGLVAGSSAERAR-----AEGLA---VHPVAE-----AAGGADIM 78
186 -ALPDVNVSHLFGKEKQSNVEVHTYRQIIVDMNPGNHLFINAVNSRDLLEIRPMP 244
79 ILDPD-----EVQRVV-----YEQETAPHLQGNVLSPFAGFN-----IHFGQIVP 119
245 GTHT-----YTLQCPALLY-----VGDSPPANDAVECNSK----- 275
120 PAHVADVNVADKGGHILVRRTYAQSEGPCLFAFYQASGARDLAMAAYAKIGSTRAGI 179
276 ----LDPKTTLL--KMAADCGGLPOISQAKLAFAFYFVGQMGYMPASM-----TR 322
180 LETFRERTETDLGEGQVVLGGGL-----SLIIRAGETVLQA-GYQBELAYFECHEVK 233
323 LMSRTASGSSVTSLDGTRSRSHTSSEGRSRSHTSSEGRSRSHTSSEGRSRSHTS 362
234 LIVDLIVEGGIAKKRDSISNTAETVAGDILTRGRIITEETRA 273

```


RESULT 14
MIG1_KLUMA STANDARD; PRT; 543 AA.
AC P52288;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Regulatory protein MIG1.
GN MIG1.
OS Kluyveromyces marxianus (Yeast) (Kluyveromyces fragilis).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=4911;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 12424;
RA Casart J., Oetling J., Rome H., Delcours J., Vandenhoute J.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INVOLVED IN GLUCOSE REPRESSION OF GLUCOSE METABOLISM
CC GENES.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE CREA/MIG GROUP OF C2H2-TYPE ZINC-
CC FINGERS PROTEINS.
CC -1- SIMILARITY: Contains 2 C2H2-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z50018; CA930321.1; -
DR HSSP; P08047; ISP2.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zF_C2H2; 2.
DR ProDom; PD000603; Znf_C2H2; 2.
DR SMART; SM00355; Znf_C2H2; 2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 2.
KW Transcription regulation; Repressor; DNA-binding; Nuclear protein;
KW Zinc-finger; Metal-binding; Repeat; Carbohydrate metabolism.
FT ZN_FING 32 54 C2H2-TYPE 1.
FT DOMAIN 60 84 C2H2-TYPE 2.
FT DOMAIN 111 131 POLY-GLN.
FT DOMAIN 499 504 POLY-SER.
SQ SEQUENCE 543 AA; 59315 MW; DFC01BD92144622P CRC64;
Query Match 4.8%; Score 99.5; DB 1; Length 543;
Best Local Similarity 19.0%; Pred No. 1.7; Indels 205; Gaps 23;
Matches 98; Conservative 53; Mismatches 159;
1 AEVKEVLKGETITGLQEPDVEODITLHGSVHVLCTGPKGNRPVILTYHDIG--M 67
3 SEVVLQKKGRKKG-----AGVVVDGKDSRPFMPCICIRGRHRL 44
68 NHKT-----CYNP--LENYEDMOEITQFAVCHVDAPG----- 98
45 EHQTSHIRHTGERPHACDFPGCAKRFGRSD--ELTRHRRIHDSKPKGRKKKSETI 102
99 -----QDDGASFPAGWY--PSMDQLAEMLPGLVQDFGL 131
103 AREKELELQRRQOOLQOQOQOOLQOQOQOVLPEIKISAPMASSMWEVTOAINQRYOD 162
132 KSIIIGMGAGAYILTRPALNPEWEGVLVINPFCAGEMWDAASKISGTOALPDMV 191
163 THNLSLSTVSGSNGSNSSNSLINVPSITOPAR-----DLA 203
192 VSHLP--GKEEMQ--SNVEVHTYRQHTVNDMPGNLHLFINA--YNSRRDLE----- 238
204 AKPMFOGSDSESESTNTTLHSVHS---QOONNGSVELLINAAFPSEKSTTLINNTPT 260

RESULT 15
DHE4_CHL50 STANDARD; PRT; 523 AA.
AC P28958;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE NADP-specific glutamate dehydrogenase (EC 1.4.1.4) (NADP-GDH)
DE (Fragment).
OS Chlorella sorokiniana.
OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
OC Chlorellaceae; Chlorella.
OX NCBI_TaxID=3076;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=92032762; PubMed=1718478;
RA Cook J.M., Kim K.D., Miller P.W., Hutson R.G., Schmidt R.R.;
RT "A nuclear gene with many introns encoding ammonium-inducible
RT chloroplastic NADP-specific glutamate dehydrogenase(s) in Chlorella
RT sorokiniana."
RL Plant Mol. Biol. 17:1023-1044(1991).
CC -1- CATALYTIC ACTIVITY: L-glutamate + H(2)O + NADP(+) = 2-oxoglutarate
CC + NH(3) + NADPH.
CC -1- SUBUNIT: HOMO- AND HETEROHEXAMER OF ALPHA AND BETA SUBUNITS. BOTH
CC SUBUNITS ARE ENCODED BY THE SAME GENE.
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- INDUCTION: BY AMMONIUM.
CC -1- PTM: THE N-TERMINUS OF THE ALPHA AND THE BETA CHAINS ARE BLOCKED.
CC -1- SIMILARITY: BELONGS TO THE GLU/LEU/PHE/VAL DEHYDROGENASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X58832; CA41636.1; -
DR EMBL; X58831; CA41635.1; ALT_SEQ.
DR FTR; S17949; S17949.
DR HSSP; P24295; LAUP.
DR InterPro; IPR006095; GLFV dehydrog.
DR InterPro; IPR006096; GLFV dehydrog. C.
DR InterPro; IPR006097; GLFV dehydrog. N.
DR Pfam; PF00208; GLFV dehydrog. 1.
DR Pfam; PF02812; GLFV dehydrog. N; 1.
DR PRINTS; PR00082; GLFV dehydrog. N.
DR PROSITE; PS00074; GLFV DEHYDROGENASE; 1.
KW Oxidoreductase; NADP; Chloroplast.
FT NON_TER 1 1
FT ACT_SITE 202 202 BY SIMILARITY.
SQ SEQUENCE 523 AA; 57529 MW; A35FE730E5FE974 CRC64;
Query Match 4.8%; Score 99; DB 1; Length 523;
239 -----IERPMGTHVTLOCPALLVVGDS-----SPA---VDAVEGNSKLDPTKT- 281
261 SFKFDLRP---PLTSSLSRALSLVSSQSPVAGAGPGGPGPANGNGANGANNGLLPRTS 317
282 -----TLIKVADCGGLPQISQPAKLAFAKFFVQ----- 310
318 RPKLSALSTLKRMTP---LSQVSEBPQCPAASLPHLQGVSSNGFLDTNGHAASVARNKSWT 374
311 GMGYMP-----SASMTRLM--RSRTASGSSVSTL----- 337
375 NLGWPMPSPSSGSGSALVSFRSSASLNKMLMDPSSRTSSAVSIATLNEDKLGQODLSV 434
435 DEFGRSRKSKSTPIRIRPSSNMGFCGPASHAME 469

Best Local Similarity 21.7%; Pred. No. 1.8;
Matches 72; Conservative 50; Mismatches 136; Indels 74; Gaps 16;

```

QY      82  MGGIITHFACVCHVDAPAGODGAASPAAGMYPSMDQLAEMPGVL-----QDFGLKSLIGM 137
Dp      225  MTELDGHISVQ--DVPAAGDIGVAAGIEGLTFQYKIRIKNTGVLTJGCGEYGSSEIRPE 283
QY      138  GTGAGAGVILTRPALNLP--EMVEGLVLI-----NV--NPAEGMDMAASKISGWTQALPD-- 189
Dp      284  ATGYGAVLFEVENVLLKDKGSELKGRKCLVSGAGNAQYCAELILKERKAVLVS-----LSDS 338
QY      190  ---MVVSHLFGKEKEMQSNVEVHTYRQHIIVDMNPNGNLHLEFINAYNSRDLRI--EREMPG 245
Dp      339  QGVYVEPNNGFTREQDQ-----YQDMKKKNNASRISEYKSDPTAVYVGDRRKPM 386
QY      246  THIVTILQCPALLVVGDSPPAVDAVVEG--NKLDTPTKTLTKMDCGGLPQ--ISQPAKLA 302
Dp      387  -----ELDLC-----QVDIAFPCCATQNIIDHDALLIKIGCCQYVEGAMPESTNE 431
QY      303  EAEFYFVQGVGYNPSPASMTLRKSRFASGVTSLDGTRSRSHTSBGRSKSHTSSECTRS 362
Dp      432  AIHKYNNAGIIYCPG-----KAANAGVAVSGLEMTQNR-----MSLMTREBYRD 477
QY      363  RSH-----TSEGAHLDTITNSGALGNSAG 386
Dp      478  KLERIMKDIYDSAMGPSREINVDLAAGAVIAG 509

```

Search completed: January 29, 2004, 23:28:24
Job time : 19 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 29, 2004, 23:25:38 ; Search time 41 Seconds
(without alignments)

2479.823 Million cell updates/sec

Title: US-10-057-832-2
Perfect score: 2072
Sequence: 1 MSRMQVDLAELVPLVKEG.....TPNSGAAGNSAGPKSMEVSC 394

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_virus: *
16: sp_bacteriap: *
17: sp_archaeap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1646.5	79.5	323	4 Q8N959	Q8N959 homo sapien
2	1227.5	59.2	388	11 Q8VCV2	Q8VCV2 mus musculu
3	1224	59.1	363	4 Q96SM2	Q96SM2 homo sapien
4	1129	54.5	339	11 Q923D7	Q923D7 mus musculu
5	1108.5	53.5	352	11 Q8BTG7	Q8BTG7 mus musculu
6	1038	50.1	371	11 Q8V101	Q8V101 rattus norv
7	1037	50.0	357	11 Q8V100	Q8V100 rattus norv
8	1036	50.0	371	11 Q8VBZ2	Q8VBZ2 rattus norv
9	1035	50.0	357	11 Q8VBW2	Q8VBW2 rattus norv
10	1023.5	49.4	356	4 Q8TRD6	Q8TRD6 homo sapien
11	948.5	45.8	306	11 Q8C6E1	Q8C6E1 mus musculu
12	891.5	43.0	253	11 Q8CB00	Q8CB00 mus musculu
13	616	29.7	485	5 Q8IG11	Q8IG11 drosophila
14	616	29.7	485	5 Q9GUS0	Q9GUS0 drosophila
15	608.5	29.4	343	5 Q8IRK7	Q8IRK7 drosophila
16	608.5	29.4	447	5 Q8T0V2	Q8T0V2 drosophila

17	601.5	29.0	343	5 Q917V6	Q917V6 drosophila
18	601.5	29.0	418	5 Q8IHC7	Q8IHC7 drosophila
19	601.5	29.0	420	5 Q8IHC2	Q8IHC2 drosophila
20	601.5	29.0	425	5 Q917V7	Q917V7 drosophila
21	601.5	29.0	446	5 Q8IHC2	Q8IHC2 drosophila
22	575.5	27.8	201	11 Q8CGR9	Q8CGR9 mus musculu
23	557	26.9	109	4 Q8BRH7	Q8BRH7 homo sapien
24	444.5	21.5	343	5 Q9U2B3	Q9U2B3 caenorhabdi
25	388.5	18.8	320	5 Q9GU49	Q9GU49 drosophila
26	362	17.5	361	5 Q8IPR2	Q8IPR2 drosophila
27	362	17.5	365	5 Q9VNH7	Q9VNH7 drosophila
28	356.5	17.2	347	10 Q9ZDN1	Q9ZDN1 arabidopsis
29	355.5	17.2	364	5 Q9ST61	Q9ST61 drosophila
30	355.5	17.2	368	5 Q9Y164	Q9Y164 drosophila
31	339.5	16.4	344	10 Q9ASU8	Q9ASU8 arabidopsis
32	338.5	16.3	346	10 Q9FJT7	Q9FJT7 arabidopsis
33	332	16.0	356	10 Q9FU21	Q9FU21 oryza sativ
34	328.5	15.8	350	10 Q9SP24	Q9SP24 helianthus
35	312	15.1	361	10 Q9LYF2	Q9LYF2 arabidopsis
36	261	12.6	132	5 Q8IH54	Q8IH54 drosophila
37	230.5	11.1	201	5 Q8IPR0	Q8IPR0 drosophila
38	153.5	7.4	63	11 Q9CTJ5	Q9CTJ5 mus musculu
39	149	7.2	335	16 Q9HZF5	Q9HZF5 pseudomonas
40	140	6.8	315	16 Q9KUG6	Q9KUG6 pseudomonas
41	138.5	6.7	275	16 Q92AY6	Q92AY6 listeria in
42	135.5	6.5	393	16 Q9ASN6	Q9ASN6 caulobacter
43	134.5	6.5	256	16 Q8X716	Q8X716 escherichia
44	131	6.3	335	16 Q8PI61	Q8PI61 xanthomonas
45	128.5	6.2	275	16 Q8Y610	Q8Y610 listeria mo

ALIGNMENTS

RESULT 1	
Q8N959	PRELIMINARY; PRT; 323 AA.
AC Q8N959;	
DT 01-OCT-2002 (TREMBLrel. 22, Created)	
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)	
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)	
DE Hypothetical protein FLJ38330.	
OS Homo sapiens (Human).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
OX NCBI_Taxid=9606;	
RN (1)	
RP SEQUENCE FROM N.A.	
RC TISSUE=Brain;	
RA Kanehori K., Ishibashi T., Chiba Y., Fujimori K., Hiraoka S.,	
RA Tanai H., Watanabe S., Ishida S., Ono Y., Hotura T., Watanabe M.,	
RA Sugiyama T., Irie R., Otsuki T., Sato H., Ota T., Wakamatsu A.,	
RA Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,	
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,	
RA Wagaetsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,	
RA Suzuki Y., Sugano S., Nagahara K., Masuno Y., Nagai K., Isogai T.,	
RT "NEO human cDNA sequencing project";	
RL Submitted (JUL-2002) to the EMBL/Genbank/DBJ databases.	
DR EMBL; AK095649; BAC04597.1; -	
DR InterPro; IPR004142; Ndr.	
DR Pfam; PF03096; Ndr; 2.	
KW Hypothetical protein.	
SK SEQUENCE 323 AA; 35226 MW; EA130D129F3B92E8 CRC64;	

Query Match 79.5%; Score 1646.5; DB 4; Length 323;
Best Local Similarity 82.0%; Pred. No. 2e-135;
Matches 323; Conservative 0; Mismatches 0; Indels 71; Gaps 1;

QY	1 MSRMQVDLAELVPLVKEGTTTGLQEPVOBODITTLGSHVVTLCGPKGRPVIL 60
DB	1 MSRMQVDLAELVPLVKEGTTTGLQEPVOBODITTLGSHVVTLCGPKGRPVIL 60
QY	61 TYHIDGMNHTCYNPLFNVEYEMOETITQHFVAVCHVDAPGQDGAASFPAQWYPSMDLAE 120

```

Db      61 TYHIDGNHKTCT----- 72
Qy      121 MLRVLQOFGKSLIIGMGTGAGATLTRPALNPNEMVEGLVINVNPAEGMMDMAASKI 180
Db      73 -----YILTRFALNNPEWEGVLINVNPAEGMMDMAASKI 109
Qy      181 SGMTQALPDVNVSHLFGKEEMQSNVEVHTYRQHI VNDMNPGNLHLFINAYSRRDLEIE 240
Db      110 SGMTQALPDVNVSHLFGKEEMQSNVEVHTYRQHI VNDMNPGNLHLFINAYSRRDLEIE 169
Qy      241 RPMFGHTVTLQCPALLVVGDSPPAVDAVVECNKLDPTKTTLLKMA DCGGLPQISOPAK 300
Db      170 RPMFGHTVTLQCPALLVVGDSPPAVDAVVECNKLDPTKTTLLKMA DCGGLPQISOPAK 229
Qy      301 LAEAFKYVQGMGMPASMTRLMRSRTASGSSVTSLDGTSRSHTSRSHTSRSHTSRGT 360
Db      230 LAEAFKYVQGMGMPASMTRLMRSRTASGSSVTSLDGTSRSHTSRSHTSRSHTSRGT 289
Qy      361 RSRSHTSRSHTSRSHTSRSHTSRSHTSRSHTSRSHTSRSHTSRSHTSRSHTSRSHTSRGT 394
Db      290 RSRSHTSRSHTSRSHTSRSHTSRSHTSRSHTSRSHTSRSHTSRSHTSRSHTSRSHTSRGT 323

```

RESULT 2

```

ID      08VCV2      PRELIMINARY;      PRT;      388 AA.
AC      08VCV2.
DT      01-MAR-2002 (TREMBLrel. 20, Created)
DT      01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT      01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE      Similar to N-myc downstream regulated 3.
GN      NDR3.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Retina;
RA      Straubeberg R.;
RL      Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR      EMBL; BC018504; AAH18504.1; -.
DR      MGI; MGI:1352499; Ndr3.
DR      InterPro; IPR004142; Ndr.
DR      InterPro; IPR000379; Ser_estrs_site.
DR      Pfam; PF03096; Ndr; 1.
SQ      SEQUENCE 388 AA; 42982 MW; AD64FE14627469C0 CRC64;

```

Query Match 59.1%; Score 1227.5; DB 11; Length 388;
 Best Local Similarity 59.1%; Pred. No. 1.1e-98;
 Matches 241; Conservative 58; Mismatches 70; Indels 39; Gaps 7;

```

Qy      4 EMQDVDLAEVPLVEKGETITGLLOEPDVOEODIETLHGSVHTLCTGPKGNRPVILTYH 63
Db      3 ELQDVQLEIKPLND-----KEHDIETHGVVHTYINGLPGKNRPVILTYH 61
Qy      64 DIGMNHKTCTNPLFNTEDMOETIOHFAVCHVDAPGQODGASFPAGTMYRSMQDLAEMLP 123
Db      62 DIGMNHKSCFNTPFNTEDMOETIOHFAVCHVDAPGQODGASFPAGTMYRSMQDLAEMLP 121
Qy      124 GVLQOFGKSLIIGMGTGAGATLTRPALNPNEMVEGLVINVNPAEGMMDMAASKISGW 183
Db      122 PVLTHLSKSLITIGVGAGATILSRFALNHPLEVEGLVLINVPCKAGMIDMAASKISGL 181
Qy      184 TQALPDVNVSHLFGKEEMQSNVEVHTYRQHI VNDMNPGNLHLFINAYSRRDLEIERPM 243
Db      182 TTNVVDIILAHNGOELQANLNDLIQTYRMHINQDINQDLQFLNSYNRDLIERPI 241
Qy      244 PGH---TYTLQCPALLVVGDSPPAVDAVVECNKLDPTKTTLLKMA DCGGLPQISOPAK 300
Db      242 LGQDNRLKTLKCTSLLVGDNSPAWEAVVECNRLDPIINTLLKMA DCGGLPQVQVPGK 301

```

```

Qy      301 LAEAFKYVQGMGMPASMTRLMRSRTASGSSVTSLDGTSRSHTSRSHTSRSHTSRSHTSRSHTSRGT 346
Db      302 LTRAFKTYFLOGMGYIPVYQVLSHSSBSPASMTRLARSHTHTSSISGSESPFSVST 361
Qy      347 SEGRSRSHTSRSHTSRSHTSRSHTSRSHTSRSHTSRSHTSRSHTSRSHTSRSHTSRSHTSRGT 394
Db      362 -----SNQSDGTQESCESP-----DVLDRH-----QTMVEVSC 388

```

RESULT 3

```

ID      096SM2      PRELIMINARY;      PRT;      363 AA.
AC      096SM2.
DT      01-DEC-2001 (TREMBLrel. 19, Created)
DT      01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT      01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE      Hypothetical protein FLJ14759.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA      Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA      Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA      Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA      Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y., Oshima A.,
RT      "MEDO human cDNA sequencing project.";
RL      Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AK027665; BAB55277.1; -.
DR      InterPro; IPR004142; Ndr.
DR      InterPro; IPR000379; Ser_estrs_site.
DR      Pfam; PF03096; Ndr; 1.
KW      Hypothetical protein.
SQ      SEQUENCE 363 AA; 39922 MW; BA1B490BE1AE78C1 CRC64;

```

Query Match 59.1%; Score 1224; DB 4; Length 363;
 Best Local Similarity 60.5%; Pred. No. 2e-98;
 Matches 239; Conservative 56; Mismatches 62; Indels 38; Gaps 6;

```

Qy      4 EMQDVDLAEVPLVEKGETITGLLOEPDVOEODIETLHGSVHTLCTGPKGNRPVILTYH 63
Db      3 ELQDVQLEIKPLND-----KEHDIETHGVVHTYINGLPGKNRPVILTYH 49
Qy      64 DIGMNHKTCTNPLFNTEDMOETIOHFAVCHVDAPGQODGASFPAGTMYRSMQDLAEMLP 123
Db      50 DIGMNHKSCNAPFNTEDMOETIOHFAVCHVDAPGQODGASFPAGTMYRSMQDLAEMLP 109
Qy      124 GVLQOFGKSLIIGMGTGAGATLTRPALNPNEMVEGLVINVNPAEGMMDMAASKISGW 183
Db      110 PVLTHLSKSLITIGVGAGATILSRFALNHPLEVEGLVLINVPCKAGMIDMAASKISGL 169
Qy      184 TQALPDVNVSHLFGKEEMQSNVEVHTYRQHI VNDMNPGNLHLFINAYSRRDLEIERPM 243
Db      170 TTNVVDIILAHNGOELQANLNDLIQTYRMHINQDINQDLQFLNSYNRDLIERPI 229
Qy      244 PGH---TYTLQCPALLVVGDSPPAVDAVVECNKLDPTKTTLLKMA DCGGLPQISOPAK 300
Db      230 LGQDNKSKTLKCTSLLVGDNSPAWEAVVECNRLDPIINTLLKMA DCGGLPQVQVPGK 289
Qy      301 LAEAFKYVQGMGMPASMTRLMRSRTASGSSVTSLDGTSRSHTSRSHTSRSHTSRSHTSRGT 359
Db      290 LTRAFKTYFLOGMGYIPVYQVLSHSSBSPASMTRLARSHTHTSSISGSESPFSVST-----SNQSDG 342
Qy      360 TRSRSHTSRSHTSRSHTSRSHTSRSHTSRSHTSRSHTSRSHTSRSHTSRSHTSRSHTSRGT 394
Db      343 TQESCESP-----DVLDRH-----QTMVEVSC 363

```

RESULT 4
 ID 0923D7 PRELIMINARY; PRT; 339 AA.

Query Match	54.5%;	Score 1129;	DB 11;	Length 339;
Best Local Similarity	63.5%;	Pred. No. 3.5e-90;		
Matches 209; Conservative	50;	Mismatches 66;	Indels 4;	Gaps 1;

RESULT 5	
Q8BTG7	
ID Q8BTG7	PRELIMINARY; PRT; 352 AA

DT 01-MAR-2003 (TREMBLER. 23, Created)
DT 01-MAR-2003 (TREMBLER. 23, last sequence update)
DT 01-MAR-2003 (TREMBLER. 23, last annotation update)
DE NDRG4 protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN
RP
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
RC MEDLINE=22354683; PubMed=12466851;
RX
RA The PANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK090374; BAC41189.1; -.
SQ SEQUENCE 352 AA; 38509 MW; 3ED388148DC852CD CRC64;

RESULT 6	
Q8VI01	
ID Q8VI01	PRELIMINARY;
	PRT; 371 AA

DT 01-MAR-2002 (TEMBLrel. 20, Created)
DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Antidepressant-related protein ADRG123.
OS Rattus norvegicus (Rat)
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Takahashi K., Yamada M., Yamada M.;
RT "Identification of a novel antidepressant related gene, ADRG123, after
RT chronic antidepressant treatment in rat frontal cortex."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF334105; AAL73186.1; --
DR InterPro:IPR004142; Ndr.
DR Pfam: PF03056; Ndr.
SQ SEQUENCE 371 AA; 40764 MW; C049A6B3F77DIDA CRC64;

[illegible]

Db 242 DLFNR-----GSEMTLKCPVNLVVGDOA.PHEDAVVEGNSKLDPTQTSTLKMADSGGQPL 297
 QY 296 SOPALAAAFKFFVQGMGYPASMTLRMRSTRASGSSVTSLDGTRSRSHSEGRSRSH 355
 Db 298 TOPGLTAAFKFFVQGMGYPASMTLRMRSTRASGSSVTSLDGTRSRSHSEGRSRSH 346
 QY 356 TSEGRSRSHSEGRSRSHSEGRSRSHSEGRSRSHSEGRSRSHSEGRSRSHSEGRSRSH 394
 Db 347 -----RTLSQSSSEGTLP-----SGPRGH-----TMEVSC 371

RESULT 7

Q8VBW2 PRELIMINARY; PRT; 357 AA.
 ID 08VBW2
 AC 08VBW2
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Antidepressant-related protein ADRG123 splice variant.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NC NCB1_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Takahashi K., Yamada M., Momose K.;
 RT "Identification of a novel antidepressant related gene, ADRG123, after
 RT chronic antidepressant treatment in rat frontal cortex."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF334106; AAL73187.1; -
 DR InterPro; IPR004142; Ndr.
 DR Pfam; PF03096; Ndr; 1.
 SQ SEQUENCE 357 AA; 39255 MW; 66792ABEA9DD9875 CRC64;

Query Match 50.0%; Score 1037; DB 11; Length 357;
 Best Local Similarity 52.4%; Pred. No. 4.1e-82;
 Matches 205; Conservative 63; Mismatches 87; Indels 36; Gaps 6;

QY 4 EMQDVDAEVRPVEKGTITGLQEPVQEDITLHGSVHTLCGTPKGNRYLTYH 63
 Db 3 ELQEVQITEKRPPLPGQTPPEA-----AKTHSVETPRGVSFTVYGTTPKRPALFTYH 55
 QY 64 DIGMNHKTCYNPLFVNEBMOEITQHFVACHVADAPQODGASFPAGYWPSPMDLAETLP 123
 Db 56 DVGANVKSQFPLPFGDMQELIQFVAVHVDAPMEBGAVPFPGYQYPSLDQADMP 115
 QY 124 GVLQOFGKLSIIGMGTGAGAYILTRFALNPMVEGLVINVNPAEGMDMAASKISGW 183
 Db 116 CILQYLNSTIIIGVAGAYILSRVYALNHPDVEGLVILNIDPAKGMMDMAHKLTL 175
 QY 184 TQALPDMVSHLFGKEBMSQVNEVHTYRQHI VDNMPNGNLHFLINAVNSRDEIERPM 243
 Db 176 TSSIEPMILGHFSGEELSGNSELIQKYSRLITTHAPNLNENTELYNNSNNRDLNFR-- 233
 QY 244 PGTHVTLOCPALVVGDSPPAVDAVEGNSKLDPTKTLTKMADCGGLPOISQPAKLA 303
 Db 234 --GSEMTLKCPVNLVVGDOA.PHEDAVVEGNSKLDPTQTSFLKMAADSGGQPLTOPGLT 291
 QY 304 AFKFFVQGMGYPASMTLRMRSTRASGSSVTSLDGTRSRSHSEGRSRSHSEGRSRSH 363
 Db 292 AFKFFVQGMGYPASMTLRMRSTRASGSSVTSLDGTRSRSHSEGRSRSHSEGRSRSH 346
 QY 364 SHTSEGAHLDTTPNSGAAGNSAGPKSMEVSC 394
 Db 336 SQSSSEGTLP-----SGPRGH-----TMEVSC 357

RESULT 8

Q8VBW2 PRELIMINARY; PRT; 371 AA.
 ID 08VBW2
 AC 08VBW2
 DT 01-MAR-2002 (TREMBlrel. 20, Created)

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE NDRG1 related protein NDRG2b1 (NDRG1 related protein NDRG2a1).
 GN NDRG2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NC NCB1_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=Sprague-Dawley; TISSUE=Heart;
 RA Boulkroun S., Fay M., Zennaro M.C., Escoubet B., Jaisser F.,
 RA Blot-Chabaud M., Farman N., Courtois-Couty N.;
 RT "Identification of rat N-myc Downstream-Regulated Gene 2 as a novel
 RT early aldosterone-induced gene."
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ26426; CAD19999.1; -
 DR EMBL; AJ26424; CAD19997.1; -
 DR InterPro; IPR004142; Ndr.
 DR Pfam; PF03096; Ndr; 1.
 SQ SEQUENCE 371 AA; 40779 MW; C049ALAAAF37D115A CRC64;

Query Match 50.0%; Score 1036; DB 11; Length 371;
 Best Local Similarity 51.6%; Pred. No. 5.4e-82;
 Matches 206; Conservative 64; Mismatches 91; Indels 38; Gaps 6;

QY 4 EMQDVDAEVRPVEKGTITGLQEPVQEDITLHGSVHTLCGTPKGN 55
 Db 3 ELQEVQITEKRPPLPGQTPPEAELAAARILID-OCQTHSVETPRGVSFTVYGTTPK 61
 QY 56 RPYLVTHYDGNMHTCYNPLFVNEBMOEITQHFVACHVADAPQODGASFPAGYWPSPM 115
 Db 62 RPAIFTHYDGNLVKSCFQPLPFGDMQELIQFVAVHVDAPMEBGAVPFPGYQYPSQ 121
 QY 116 DQLAEMLPGLVQOFGKLSIIGMGTGAGAYILTRFALNPMVEGLVINVNPAEGMDW 175
 Db 122 DQADLMPICILQYLNSTIIIGVAGAYILSRVYALNHPDVEGLVILNIDPAKGMMDW 181
 QY 176 AASKISGWTQALPDMVSHLFGKEBMSQVNEVHTYRQHI VDNMPNGNLHFLINAVNSR 235
 Db 182 AAHLTGLTSSIEPMILGHFSGEELSGNSELIQKYSRLITTHAPNLNENTELYNNSNNR 241
 QY 236 DLBIERPMPGTHVTLOCPALVVGDSPPAVDAVEGNSKLDPTKTLTKMADCGGLPQI 295
 Db 242 DLFNR-----GSEMTLKCPVNLVVGDOA.PHEDAVVEGNSKLDPTQTSFLKMAADSGGQPL 297
 QY 296 SOPALAAAFKFFVQGMGYPASMTLRMRSTRASGSSVTSLDGTRSRSHSEGRSRSH 355
 Db 298 TOPGLTAAFKFFVQGMGYPASMTLRMRSTRASGSSVTSLDGTRSRSHSEGRSRSH 346
 QY 356 TSEGRSRSHSEGRSRSHSEGRSRSHSEGRSRSHSEGRSRSHSEGRSRSHSEGRSRSH 394
 Db 347 -----RTLSQSSSEGTLP-----SGPRGH-----TMEVSC 371

RESULT 9

Q8VBW2 PRELIMINARY; PRT; 357 AA.
 ID 08VBW2
 AC 08VBW2
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE NDRG1 related protein NDRG2b2 (NDRG1 related protein NDRG2a2).
 GN NDRG2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NC NCB1_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=Sprague-Dawley; TISSUE=Heart;
 RA Boulkroun S., Fay M., Zennaro M.C., Escoubet B., Jaisser F.,
 RA Blot-Chabaud M., Farman N., Courtois-Couty N.;

"Identification of rat N-myc Downstream-Regulated Gene 2 as a novel
RT easily aldosterone-induced gene."
RL Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AJ426427; CAD0000.1; -
DR EMBL: AJ426425; CAD0000.1; -
DR InterPro: IPR004142; Ndr.
DR Pfam: PF03096; Ndr; 1.
SQ SEQUENCE 357 AA; 39270 MW; 66729DC7A9DD9485 CRC64;

Query Match 50.0%; Score 1035; DB 11; Length 357;
Best Local Similarity 52.4%; Pred. No. 6.2e-82;
Matches 205; Conservative 62; Mismatches 88; Indels 36; Gaps 6;

QY 4 EMODVLAEPVLEKEGTITGLLOEPVQODIETLHGSVHTLCGTPKGNRPVILTYH 63
DB 3 ELQVOITBEKRL-PCGTPEA-----AKHSVETPGSVTFYVGGPKRPAIFTYH 55
QY 64 DIGMNHKTCTNPLFNYEDMOEITQHFVAVCHVDAPQODGAASFPAGYWPMDQLAEMLP 123
DB 56 DVLGNYKSCFQPLQFQDMQEIIGNFVAVHVDAPGMEGAVPFLGYQPSQDQLADWIP 115
QY 124 GVLDQFGIAKSIIGMGTGAGAYILTRFALNPEMVEGLIVNVPACBGMMDMAASKISGM 183
DB 116 CILDYLNSTIIGVGAGAYILSRVALNHPDVEGLVINDPAKGMMDMAAKLTGL 175
QY 184 TQALPDVAVSHLFGKEEQSNVEVHTYRQHVNDMPGNLHLFINAYSRRDLEIERPM 243
DB 176 TSSIPEMILGHLFSGEELSGNSELIQKTRSLTHAPNLNLELYNNSYNNRRDLNFER-- 233
QY 244 PGHTVTLQCPALLVVDSSPAVDVAVECNKLDPTKTLTKMADCGGLPOISOPAKLAE 303
DB 234 --GGEMTLKCPVMLVVGQAHPEDAIVVECNKLDPTQTSFLKMDSGGQPOLQPGKLT 291
QY 304 AFKTFVQGMGMPASMTRLMRSTRASGVTSLDGTSRSRHTSGTSRSRHTSGTSR 363
DB 292 AFKTFVQGMGMASSCMTRLSRSTRASJLTSASIDGSRSS-----RTL 335
QY 364 SHTSEGAHLDTTPNSGAAGNSAGPKSMEVSC 394
DB 336 SSSSESGTLP-----SGPPGH-----TMEVSC 357

RESULT 10

Q8TRD6 PRELIMINARY; PRT; 356 AA.

AC Q8TRD6 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Cytoplasmic protein Ndr1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhou Y., Yu L., Zhao Y., Fan Y.X., Huang J., Zhao S.Y.;
RT "Cloning and expression of a novel human cDNA homology to murine
cytoplasmic protein Ndr1 mRNA."
RL Submitted (Apr-2002) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yue P., Yu L., Ding J.B., Zhao Y., Huang J., Zhao S.Y.;
RL Submitted (Aug-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF087872; AAM10500.1; -
DR InterPro: IPR004142; Ndr.
DR Pfam: PF03096; Ndr; 1.
SQ SEQUENCE 356 AA; 39531 MW; E9470B6A701B9C17 CRC64;

Query Match 49.4%; Score 1023.5; DB 4; Length 356;
Best Local Similarity 55.3%; Pred. No. 6.2e-81;
Matches 198; Conservative 56; Mismatches 75; Indels 29; Gaps 4;

QY 37 IETLHGSVHTLCGTPKGNRPVILTYHDIGMNHKTCTNPLFNYEDMOEITQHFVAVCHVDA 96
DB 28 VETPEGVTFTVYCTPKRPAIFTYDVLGNYKSCFQPLQFQDMQEIIGNFVAVHVD 87
QY 97 PGQODGAASFPAGYWPMDQLAEMLPVGLQFGLKSIIGMGTGAGAYILTRFALNPEM 156
DB 88 PGMEGAPVPEPLQGYPSLDQADMIPCVLOYLNFSITIGVAGAYILARVALNHPDT 147
QY 157 VEGVILNVPACBGMMDMAASKISGWTQALPDVAVSHLFGKEEQSNVEVHTYRQHV 216
DB 148 VEGVILNVPACBGMMDMAAKLTGLTSSIPEMILGHLFSGEELSGNSELIQKTRNIT 207
QY 217 NDMPGNLHLFINAYSRRDLEIERPMGHTVTLQCPALLVVDSSPAVDVAVECNK 276
DB 208 HAPLNDIELYNNNSYNNRRDLNFER-----GGDITLRCPVMLVVGQAHPEDAIVVECNK 263
QY 277 DPTKTLTKMADCGGLPOISOPAKLAEAFKTFVQGMGMPASMTRLMRSTRASGVTS 336
DB 264 DPTQTSFLKMDSGGQPOLQPGKLTLEAFKTFVQGMGMASSCMTRLSRSTRASJLTSAS 323
QY 337 LDGTSRSRHTSGTSRSRHTSGTSRSRHTSGTSRSRHTSGTSRSRHTSGTSRSRHTSGTSR 394
DB 324 VDGNRSS-----RTLSQSSSESG-----TSSGPPGH-----TMEVSC 356

RESULT 11

Q8C661 PRELIMINARY; PRT; 306 AA.

AC Q8C661 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE N-myc downstream regulated 2 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL: AK076514; BAC36373.1; -
DR NON TER 1
FT 1
SQ SEQUENCE 306 AA; 33701 MW; 42B01C152E197D2F CRC64;

Query Match 45.8%; Score 948.5; DB 11; Length 306;
Best Local Similarity 55.4%; Pred. No. 1.8e-74;
Matches 185; Conservative 53; Mismatches 67; Indels 29; Gaps 5;

QY 61 TYHDIGMNHKTCTNPLFNYEDMOEITQHFVAVCHVDAPQODGAASFPAGYWPMDQLAE 120
DB 2 TYHVGILNYKSCFQPLQFQDMQEIIGNFVAVHVDAPGMEGAVPFLGYQPSLDQAD 61
QY 121 MLPVGLQFGLKSIIGMGTGAGAYILTRFALNPEMVEGLIVNVPACBGMMDMAASKI 180
DB 62 MIPCLQILNFSITIGVGAGAYILSRVALNHPDVEGLVINDPAKGMMDMAAKLT 121
QY 181 SGWTQALPDVAVSHLFGKEEQSNVEVHTYRQHVNDMPGNLHLFINAYSRRDLEIE 240
DB 122 TGLTSSIPDMLGHLFSGEELSGNSELIQKTRGIIQHPNLNLELYNNSYNNRRDLNFB 181
QY 241 RPMGHTVTLQCPALLVVDSSPAVDVAVECNKLDPTKTLTKMADCGGLPOISOPAK 300
DB 182 R---GERT-TLKCPVMLVVGQAHPEDAIVVECNKLDPTQTSFLKMDSGGQPOLQPGK 237
QY 301 LAEPKTFVQGMGMPASMTRLMRSTRASGVTSLDGTSRSRHTSGTSRSRHTSGTSR 360
DB 238 LTEAFKTFVQGMGMASSCMTRLSRSTRASJLTSASIDGSRSS----- 281

Qy 361 RSRHTSEGALHDITPNSGAGNAGSAPKSMVEVSC 394
 Db 282 RLSQSSSESGLT-----SGPCH-----TMEVSC 306

RESULT 12

Q8CBDO PRELIMINARY; PRT; 253 AA.

Qy 08CBDO: 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE N-myc downstream regulated 3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RP SEQUENCE FROM N.A.
 RX STRAIN=C57BL/6J; Tissue=Cerebellum;
 MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 DR EMBL; AK036227; BAC29369.1; -.
 SO SEQUENCE 253 AA; 28583 MW; 568DCTF547BDF0A3 CRC64;

Query Match 43.0%; Score 891.5; DB 11; Length 253;
 Best Local Similarity 66.1%; Pred. No.1.3e-69;
 Matches 160; Conservative 40; Mismatches 41; Indels 1; Gaps 1;

Qy 4 EMOVDLAEVKPLEKGETITGLLOEPVODIETLHGSVHTLCGTPKGNRPVILTYH 63
 Db 3 ELQDVLLEIKRPL-NDKNGTRNFPQDCEHDIETPHGMVHTIRGLPKGNRPVILTYH 61
 Qy 64 DIGMNHKTCTNPLPYEDMOEITQFAVCHVDAPQODGAASFPAGWYPSMDLAEMLP 123
 Db 62 DIGMNHKSCFTFFNFEDMOEITQFAVCHVDAPQODGAASFPAGWYPTMDLAEMLP 121
 Qy 124 GVLQDFGLKSIIGMGTGAGAYILTRFALNPEMVEGLVINVPACBGMMAASKISGW 183
 Db 122 PVLFTLSMKSIIGIGVGAGAYILSRFALNHPDELVEGLVINIDPCAKIMMAASKISGF 181
 Qy 184 TQALPDWVSHLFGKEEWSNVEVHTYRQHIIVNDMPGNLHLFINAYNSRDLIERPM 243
 Db 182 TTNIIVDITLHFGQELQANMLDIQYRLHADIQINQENLQPLFGSYNGRRDLIERPI 241
 Qy 244 PG 245
 Db 242 LG 243

RESULT 13

Q8IGI1 PRELIMINARY; PRT; 468 AA.

Qy 08IGI1: 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE RH13074p.
 GN MESK2.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RP SEQUENCE FROM N.A.
 RC STRAIN=y;
 RA Stapleton M., Brokstein P., Hong L., Agdayani A., Carlson J.,

RA Chame M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guatin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Ceiniker S.,
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BT001773; AAN71528.1; -.
 SO SEQUENCE 468 AA; 51588 MW; E3CDAL48332DDCC3 CRC64;

Query Match 29.7%; Score 616; DB 5; Length 468;
 Best Local Similarity 35.0%; Pred. No.3.8e-45;
 Matches 151; Conservative 62; Mismatches 162; Indels 56; Gaps 11;

Qy 5 MODVDLAEVKPLEKGETITGLLOEPVODIETLHGSVHTLCGTPKGNRPVILTYH 64
 Db 6 MDDELASVGLQPPNAR---GSILF-ACEGRRPVTDGVDHVAIQGDTA--KRAIITYHD 59
 Qy 65 IGMMHKTCTNPLPYEDMOEITQFAVCHVDAPQODGAASFPAGWYPSMDLAEMLP 124
 Db 60 LGINVATSFAGFFNFPVWRGLLENFCVYHTARQEGEAPTLPEDDVYPTMDLAQLLF 119
 Qy 125 VLOQDFGLKSIIGMGTGAGAYILTRFALNPEMVEGLVINVPACBGMMAASKISGW 178
 Db 120 VLSHFGLKSVIGGVGAGANIILARFALHAPDKXGALCLINCVSTQSGMIEMGQSFNARF 179
 Qy 179 -KISGWTQALPDWVSHLFGKEEWSNVEVHTYRQHIIVNDMPGNLHLFINAYNSRDL 237
 Db 180 LRTKGMTQSVITDLMHMFGRNBERNHDVQWTKQHFEGVAPPTLAMIINYYIRNDL 239
 Qy 238 EIERPMFG-----THVTTLQCPALLVGDSSPAVDAVVEGNSKLDPTKTLKPADCGGL 292
 Db 240 HLRATPGTGSFAATTLKMPVINITGSLSPHYDDVTFTNGRLDPTNSMMKISDC-AL 298
 Qy 293 PQISQPAKLAERKYPVQNGY-----MPSAS-----MTRL 323
 Db 299 VLEBQPKLAEARFLFLOGGGYATPLSTPASSCGCTKYHTYSSIFPANFREDOQAHEER 358
 Qy 324 MRSRTASGSSVTSILD---GTRSRHTSEGRSRHTSEGRSRHTS---EGALHDITP 376
 Db 359 DREPERRDQRQLSLVRGRLRETAINCTSSNSNTTAAAGGDNNMADGRVDDDDQLLEN 418
 Qy 377 NSG-AAGNSAG 386
 Db 419 GNGCAGGNGTG 429

RESULT 14

Q9GU50 PRELIMINARY; PRT; 485 AA.

Qy 09GU50: 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Msexpression suppressor of KSR (GH09802p).
 GN MESK2 OR CG15668 OR CG15669.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RP SEQUENCE FROM N.A.
 RA Huang A.M., Rubin G.M., Tsang G., Evans-Holm M., Suh C.S.,
 RT "Full length Drosophila melanogaster cDNA sequence."
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 RC STRAIN=y;
 RA Stapleton M., Brokstein P., Hong L., Agdayani A., Carlson J.,
 RA Chame M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guatin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,

RA Ceiniker S.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP195792; AAC82487.1; -
 DR EMBL: BT001407; AAN71152.1; -
 DR FlyBase; FBgn043070; MESK2.
 DR InterPro; IPR004142; Ndr.
 DR Pfam; PF03096; Ndr; 1.
 DR SEQUENCE 485 AA; 53333 MW; 55C19C656736FA6C CRC64;
 SO
 Query Match 29.7%; Score 616; DB 5; Length 485;
 Best Local Similarity 35.0%; Pred. No. 4e-45;
 Matches 151; Conservative 62; Mismatches 162; Indels 56; Gaps 11;
 5 MODVLAEVKVLVEKGETITGLQEPVQEDIEHLSGSHVHTLCGTPKGRNPVILTYHD 64
 23 MDDIELSLVQLQFNAR---GSILE-ACEQRRVPTDKADVAALQGDTA--KPAITTHD 76
 65 IGNNKTCYNPLFVYEDMQEITQHFVAVCHVADAPGQDGAASFPAGYWYPSMDLAEMLP 124
 77 LGLVYATSPAFGFPFVVRGLENVCVHTAPQGBEAPTLPEYVYPTMDLLAQLLF 136
 125 VLOQGLKSIIGMTGAGVILTRPALNPEWBGVLINVPACBGMWMAAS----- 178
 137 VLSHFGLKSVIGFVGAGANILARFAHAPDKVGLCLINCSTQSGWIEWYOSFNARF 196
 179 -KISGMQALPDMVSHLFGKEEMQSNVEVHTYRQHTVNDMNQNLTLFINAYNSRDL 237
 197 LRTGTMGTGVIDVLMHMFHGNPERNHDLVOMKQHERGVNPTNLMLINATYHRDL 256
 238 EIERPMFG-----THVTIQCPLALIVGSSPAVDAVECSKLDPTKTLTKMADCGGL 292
 257 HLAITPPTPSSERFATTLKRPVINITSGLSPHDDVYTFNGRLDPTNSSMVKISDC-AL 315
 293 PQLSQPAKLAFAFKYFVQGMG---MPSAS-----MTRL 323
 316 VLEQPAKLAFAFLFLQEGEATPLSTPASPCTKYHTYSSIFPANFRECOQOAMBER 375
 324 MRSRTASSSVTSYSD---GTSRSRHSRTSRTSRTSRTSRTSRTSRTSRTSRTSRTS 376
 376 DREBRERDRORQLSLRVGNRLRETAINCTSSNTTAGQGDNNADRSVGDGDDQLEN 435
 377 NSG-AAGNSAG 386
 436 GNGCAGCAGNGTG 446
 RESULT 15
 OSIRK7 PRELIMINARY; PRT: 343 AA.
 ID OSIRK7
 AC OSIRK7
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE CG15669-PD.
 GN MESK2.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephyridiidae; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoekins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang O., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazer R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA Abill J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailew R.W., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz S., Ferreira S., Fleischmann W.,
 RA Foeller C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Gloder A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Hardin N.L., Hatvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibbegam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., Mckerson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Pacled J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler T., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wassarman D.A., Weinrock G.M., Weissbach J.,
 RA Williams S.M., Woodagel, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers B.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ceiniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Barzon J., An H., Baldwin D., Barzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champagne M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
 RA Ferris J., Frise B., Galie R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoekins R.A., Hostin D., Howland T.J.,
 RA Ibbegam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Pacled J., Paragas V., Park S., Patel S., Pfeiffer B., Scheeler F.,
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of Drosophila melanogaster genome."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hiradecky P., Huang Y., Kaminker J.S., Prochuk S.B., Smith C.D.,
 RA Tupy J.L., Bergman C., Berman C., Carlson J.W., Ceiniker S.E.,
 RA Clump M., Drysdale R., Emert D., Frise B., de Grey A., Harris N.,
 RA Kronmiller B., Marshall B., Milburn G., Richter J., Russo S.,
 RA Seearle S.M.J., Smith E., Shu S., Smutnialak F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
 RT "Annotation of Drosophila melanogaster genome."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Adams M.D., Ceiniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003454; AAM70904.2; -
 DR SEQUENCE 343 AA; 37936 MW; AF87CF7B61157A67 CRC64;
 SO
 Query Match 29.4%; Score 608.5; DB 5; Length 343;
 Best Local Similarity 38.8%; Pred. No. 1.1e-44;
 Matches 135; Conservative 57; Mismatches 13; Indels 19; Gaps 6;
 5 MODVLAEVKVLVEKGETITGLQEPVQEDIEHLSGSHVHTLCGTPKGRNPVILTYHD 64

```

Db      1 MDIEILRSVQLQFPNAR---GSILE-ACEQRRVPTDKGDVHVAIQGDTA--KPAIITYHD 54
QY      65 IGMNHTCYNPLFNYEDMOEITTOHFAVCHVDAFGODGAASFPAQVYPPSMDOLAEMLPG 124
Db      55 LGLNVAISFAGFNFVVRGLENFCYHVTAPQGBGAPTLPEYVYPTMDLAAQDLF 114
QY      125 VLOQFGKSIIGMGTGAGAYILTRFALNNPEMEVGLVLINVPCEAGMMDWAAS----- 178
Db      115 VLSHFGKSVIGFVGAGANILARFAHAHPDKVGALCLINCSTQSGWIEKGYQSFNARF 174
QY      179 -KISGWTQALPDVNVVSHLFGKEEMQSNVEVVTYRQHI VDNMPGNLHLFTNAYNSRRDL 237
Db      175 LRTGWTQGVLDYLMHHFGRNPEERHNDLVQMYTKQHFBERGVNPTNLAMLINAYIHRNDL 234
QY      238 EIERPMPG-----THVTLOCPALLVWGDSSPAYDAVEECNSKLDPTKTLTKMADCGGL 292
Db      235 HLAHTPPGTGSETAATTLKMPVINITGSLSPHYDDTVTFNGRLDPTNSSMWKISDC-AL 293
QY      293 POISQPAKLAFAFKYFVQGMGMPSPASWTRLMRSRTASGSSVTSIDGT 340
Db      294 VLEQPAKLAFAFRLFLQGBGYAVGTLOKLARKISSVSRSSSTQIELT 341

```

Search completed: January 29, 2004, 23:29:18
 Job time : 43 secs